



**FACULTAD DE CIENCIAS EMPRESARIALES
MAESTRÍA EN ADMINISTRACIÓN ECONÓMICA Y FINANCIERA**

**MODELO SCORING PARA OTORGAMIENTO DE CRÉDITO COMERCIAL
EN UNA ENTIDAD FINANCIERA A INGENIEROS CIVILES CONTRATISTAS**

**Tesis de Maestría
Karina Alzate Quintero
Paulo César Mesa Martínez**

Pereira, 2018

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
FACULTAD DE CIENCIAS EMPRESARIALES
MAESTRÍA EN ADMINISTRACIÓN ECONÓMICA Y FINANCIERA

MODELO SCORING PARA OTORGAMIENTO DE CRÉDITO COMERCIAL
EN UNA ENTIDAD FINANCIERA A INGENIEROS CIVILES CONTRATISTAS

Tesis de Maestría realizada por:

Karina Alzate Quintero, Código: 41'942.563

Paulo César Mesa Martínez, Código 89'000.847

Bajo la dirección y el asesoramiento del profesor

M.Sc. Leonel Arías Montoya

Pereira, 2018

Índice general

1	INTRODUCCIÓN	7
2	EL PROBLEMA DE INVESTIGACIÓN	9
2.1	Antecedentes de la idea	9
2.2	Situación problema	12
2.3	Definición del problema	13
2.4	Hipótesis o supuestos	13
2.5	Objetivo general	13
2.5.1	Objetivos específicos	13
2.6	Justificación del estudio	14
2.7	Beneficios que conlleva	15
2.8	Limitaciones previsibles	15
2.9	Metodología	16
3	MARCO REFERENCIAL	17
3.1	Marco Teórico	17
3.1.1	El Sistema Financiero Colombiano	17
3.1.2	Modelos scoring	23
3.1.3	Modalidades de contratación pública	26
3.2	Marco Conceptual	32
3.2.1	Riesgo de crédito	32
3.2.2	Introducción al Análisis Discriminante	35
3.3	Marco Normativo	45
3.3.1	Los Acuerdos de Basilea	45
3.3.2	Normatividad que rige a la población objeto de estudio	48

3.4	Marco Filosófico	48
3.5	Marco Situacional	49
4	DISEÑO METODOLÓGICO	51
4.1	Introducción	51
4.2	Población o universo	52
4.3	Muestra	52
4.4	Delimitación del estudio	52
4.5	Variables e indicadores del estudio	53
4.6	Instrumentos para recolección de la información	56
4.7	Procesamiento y análisis de la información	56
5	MODELO DE CREDIT SCORING PARA EL SICC	57
5.1	Modelación por medio de análisis discriminante	57
5.1.1	Uso de Infostat	58
5.1.2	Variables	59
5.2	Modelación	61
5.2.1	Análisis descriptivo previo	61
5.2.2	Variables para el análisis discriminante	76
5.3	Proceso de modelación	76
5.3.1	Interpretación de resultados	80
5.3.2	Validación: Back Testing	81
5.3.3	Análisis cualitativo de resultados	83
6	CONCLUSIONES	84
7	BIBLIOGRAFÍA	88

Índice de cuadros

Cuadro4.1	Variables a considerar en el modelado (1).	54
Cuadro4.2	Variables a considerar en el modelado (2).	55
Cuadro5.1	Variables de la base de datos inicial.	60
Cuadro5.2	Estadísticos Chi-cuadrado de Pearson y MV-G2	76
Cuadro5.3	Variables a tener en cuenta en la aplicación del Análisis discriminante	77
Cuadro5.4	Back testing de la validación del modelo	82

Índice de figuras

Figura3.1	Billete de un peso colombiano.	18
Figura3.2	Punto de corte discriminante	45
Figura5.1	Logo del software estadístico Infostat	58
Figura5.2	Ruta en Infostat para realizar análisis discriminante	58
Figura5.3	Ingreso de las variables	59
Figura5.4	Estadísticas de la variable <i>Número de contratos adjudicados</i>	62
Figura5.5	Estadísticas discriminadas según grupo para la variable <i>Número de contratos adjudicados</i>	62
Figura5.6	Análisis de varianza en la variable <i>Número de Contratos Adjudicados</i>	63
Figura5.7	Estadísticas de la variable <i>Edad</i>	63
Figura5.8	Estadísticas discriminadas según grupo para la variable <i>Edad</i>	63
Figura5.9	Estadísticas de la variable <i>Hijos</i>	64
Figura5.10	Estadísticas discriminadas según grupo para la variable <i>Hijos</i>	64
Figura5.11	Estadísticas para la variable <i>Ingresos anuales</i>	64
Figura5.12	Estadísticas discriminadas según grupo para la variable <i>Ingresos anuales</i>	65
Figura5.13	Análisis de varianza para la variable <i>Ingresos Anuales</i>	65
Figura5.14	Estadísticas de la variable <i>Número de entidades contratantes</i>	65
Figura5.15	Estadísticas discriminadas según grupo para la variable <i>Número de entidades contratantes</i>	66
Figura5.16	Análisis de varianza para la variable <i>Número de entidades contratantes</i>	66
Figura5.17	Estadísticas para la variable <i>Número de años de experiencia</i>	66
Figura5.18	Estadísticas según grupo para la variable <i>Número de años de experiencia</i>	67
Figura5.19	Estadísticas para la variable <i>Número de personas a cargo</i>	67
Figura5.20	Estadísticas según grupo para la variable <i>Número de personas a cargo</i>	67

Figura5.21 Estadísticas para la variables: <i>Garantía, Nivel de endeudamiento, Género, Estrato, Estado civil, Tipo de vivienda, Otros ingresos y Calificación Cifin.</i>	69
Figura5.22 Variable ESTRATO renombrada como CATESTRATO.	71
Figura5.23 Estadísticas variable ESTRATO.	71
Figura5.24 Tabla de contingencia variables CIF IN vs PAGODEFAULT	72
Figura5.25 Tabla de contingencia variables GENERO vs PAGODEFAULT	72
Figura5.26 Tabla de contingencia variables GARANTIA vs PAGODEFAULT	73
Figura5.27 Tabla de contingencia variables VIVIENDA vs PAGODEFAULT	73
Figura5.28 Tabla de contingencia variables ESTADO_CIVIL vs PAGODEFAULT	74
Figura5.29 Tabla de contingencia variables OTROS_INGRESOS vs PAGODEFAULT	74
Figura5.30 Tabla de contingencia variables NE vs PAGODEFAULT	75
Figura5.31 Ingreso de variables a las aplicación Infostat para modelado	78
Figura5.32 Opciones escogidas para el análisis	78
Figura5.33 Salida del programa Infostat para análisis discriminante (I).	79
Figura5.34 Salida del programa Infostat para análisis discriminante (II).	79

Dedicamos éste trabajo a nuestras familias, a los docentes que durante éste posgrado nos compartieron sus conocimientos, a las directivas de la Maestría en Administración Económica y Financiera y en general a la Universidad Tecnológica de Pereira

Karina y Paulo César

Agradecimientos

A mi esposo Jhon Jairo,
A mis hijos Cristian y Juan Daniel,
A mi madre Martha Isabel.

Karina

Agradezco profundamente a Dios
y a mi familia: Luz Karime, Lucia, Violeta y Julián.

Paulo César

Resumen

El presente trabajo cumplió el objetivo de *Diseñar un modelo scoring para otorgamiento de crédito comercial a ingenieros civiles contratistas en una entidad financiera del municipio de Armenia.*

El trabajo se centró en el análisis de una base de datos suministrada por una entidad financiera de la ciudad de Armenia la cual solicitó reservar su nombre y por tal razón en éste trabajo fué llamada **EFArmenia**. Esta base de datos cuenta con un total de 101 clientes del Segmento de Ingenieros Civiles Contratistas (SICC) que han adquirido créditos con la entidad y de los cuales el 58,23% han cumplido con sus pagos oportunamente. Hay un 41,77% que se encuentran en la categoría de morosos debido a que han tenido algún problema de cumplimiento con sus pagos. La base de datos acumula un total de 15 variables para cada cliente, entre cualitativas y cuantitativas. El SICC de categoría PYME, es un segmento bastante particular, con unas características propias que sustentaron la justificación de este trabajo en cuanto a crear un modelo scoring debido a que los modelos scoring actuales utilizados en el otorgamiento de crédito a éste tipo de PYMES no son acertados en la medición de probabilidad de incumplimiento. El 41,77% de clientes que incumplieron en sus pagos fueron aprobados por medio de los análisis tradicionales para PYMES. Esta cifra es muy alta comparada con el índice de cartera comercial vencida para los segmentos PYMES del 7% y el índice de cartera comercial vencida para el SICC del 27% con corte al 31 de marzo de 2017 en la región sur de la EFArmenia¹.

El diseño del modelo propuesto se hizo utilizando la técnica de análisis discriminante sobre el software Infostat. Para el ejercicio se utilizaron 79 datos y se dejaron 22 para su validación. Se obtuvo así un modelo scoring para el otorgamiento de crédito comercial

¹Fuente: EFArmenia.

al SICC, el cual comparado con el convencional que utiliza la entidad financiera EF Armenia presenta una mayor eficiencia, en cuanto al contrastarlo con la base de datos hace una discriminación de los clientes con un 99,9% de efectividad clasificando los clientes en default (no pago) con un error del 0,00% y los clientes en pago (no morosos) con un error de solo el 0,045%. El modelo scoring final cuenta con 5 variables discriminatorias, a saber: Número de contratos adjudicados (x_1), Ingresos anuales (x_2), Número de entidades contratantes (x_3), Otros ingresos (x_4) y Tipo de vivienda (x_5).

El modelo scoring resultante fué:

$$D = -2,52 + 0,07x_1 + 6,2 \times 10^{-11}x_2 + 0,93x_3 + 0,55x_4 - 0,55x_5.$$

La conclusión principal es que la EF Armenia debe manejar éste modelo como complemento para el análisis de riesgo en el otorgamiento de crédito. La propuesta es utilizar éste modelo luego de aplicar el análisis de cifras financieras que actualmente utiliza la EF Armenia (RiskAnalyst). Se prevee que la utilización del modelo para la medición de la probabilidad de incumplimiento va a mejorar en un gran porcentaje en relación al proceso tradicional utilizado. Queda abierta la posibilidad de evaluar otros segmentos que posiblemente ameriten un modelo particular para la medición de la probabilidad de incumplimiento.

Abstract

The present work fulfilled the objective of *Designing a scoring model for granting commercial credit to civil engineers contractors in a financial entity of the municipality of Armenia*.

The work focused on the analysis of a database provided by a financial institution of the city of Armenia which requested to reserve its name and for this reason in this work was called **EFArmenia**. This database has a total of 101 clients of the Segment of Civil Engineers Contractors (SICC) that have acquired credits with the entity and of which the 58,23% have fulfilled their payments in a timely manner. There is a 41,77% that is in the category of defaulters because they have had some problem of compliance with their payments. The database accumulates a total of 15 variables for each client, between qualitative and quantitative. The SICC of SME category, is a very particular segment, with its own characteristics that supported the justification of this work in terms of creating a scoring model because the current scoring models used in the granting of credit to this type of PYMES are not successful in measuring the probability of default. The 41,77% of clients who failed to make their payments were approved through traditional analysis for PYMES. This figure is very high compared to the commercial loan portfolio index for the PYMES segments of 7% and the commercial loan portfolio index for the SICC of 27% with a cut as of March 31, 2017 in the region south of the EFArmenia².

The design of the proposed model was done using the discriminant analysis technique on the Infostat software. For the exercise, 79 data were used and 22 were left for validation. A scoring model was thus obtained for the granting of commercial credit to the SICC, which, compared to the conventional one used by the financial institu-

²Source: EFArmenia.

tion EFArmenia, shows greater efficiency, in contrast to the database it discriminates against customers with a 99,9% of effectiveness classifying the clients in default (non-payment) with an error of 0,00% and customers in payment (not delinquent) with an error of only 0,045%. The final scoring model has 5 discriminatory variables, namely: Number of contracts awarded (x_1), Annual revenues (x_2), Number of contracting entities (x_3), Other income (x_4) and Type of housing (x_5).

The resulting scoring model was:

$$D = -2,52 + 0,07x_1 + 6,2 \times 10^{-11}x_2 + 0,93x_3 + 0,55x_4 - 0,55x_5.$$

The main conclusion is that the EFArmenia should handle this model as a complement to the risk analysis in the granting of credit. The proposal is to use this model after applying the analysis of financial figures currently used by the EFArmenia (RiskAnalyst). It is expected that the use of the model for the measurement of the probability of non-compliance will improve in a large percentage in relation to the traditional process used. The possibility of evaluating other segments that possibly merit a particular model for the measurement of the probability of non-compliance is still open.

INTRODUCCIÓN

En el presente trabajo se propone un modelo scoring como complemento al proceso de otorgamiento de crédito en una entidad financiera de la ciudad de Armenia al Segmento de Ingenieros Civiles Contratistas (SICC) que hasta el momento se ha realizado con los modelos convencionales para PYMES, situación que ha generado un alto nivel de cartera vencida comparado con indicadores de otros segmentos similares de actividades económicas diferentes. Se detecta en este segmento unas particularidades que los hacen diferentes. Una de estas particularidades es la concentración de ingresos para un empresa de este segmento que en muchos casos está por encima del 60% en un mismo cliente o entidad contratante, mientras que el mismo volumen de ingresos para un empresa PYME de otro sector puede estar referido a una gran cantidad de clientes. Este indicador medido desde el punto de vista financiero podría para ambos casos tener un resultado similar en la evaluación de crédito. En contraste, los resultados de cartera morosa no son iguales, siendo mucho más alto en el segmento estudiado. Según la entidad financiera en estudio, que en adelante se llamará en éste trabajo EFArmenia, el 41,77% de clientes que incumplieron en sus pagos fueron aprobados por medio de

los análisis tradicionales para PYMES. Esta cifra es muy alta comparada con el índice de cartera comercial vencida para los segmentos PYMES del 7% y el índice de cartera comercial vencida para el SICC del 27% con corte al 31 de marzo de 2017¹.

En éste trabajo de investigación se seleccionaron únicamente los ingenieros civiles contratistas persona natural que contratan con el estado bajo la reglamentación del Estatuto General de Contratación de la Administración Pública (Ley 80 de 1993 y la Ley 1150 de 2007), clientes de una entidad financiera de la ciudad de Armenia que de acuerdo a los servicios prestados celebran contratos bajo la modalidad de contrato de obra, para la construcción especialmente de obras de ingeniería como acueducto, alcantarillado, pavimentos, vías, puentes y todo lo relacionado con obra civil, independientemente de la modalidad de la contratación, sin embargo para este tipo de contrato de obra las más comunes son licitación pública y selección abreviada, y sin importar cuál sea la modalidad de ejecución y pago, estos clientes cumplen también la característica de ser segmentados como clientes PYME, que según política de la entidad financiera sus ingresos anuales deben ser superiores a 400 millones de pesos.

Este trabajo está estructurado en 6 capítulos. El primero de ellos es éste capítulo introductorio. El Capítulo 2, presenta el problema de investigación. El Capítulo 3, hace una presentación del marco referencial. El Capítulo 4, hace referencia al diseño metodológico. El Capítulo 5, presenta el modelo scoring para el SICC. En él se hace desde una presentación teórica de la técnica del análisis discriminante, hasta la aplicación de la misma con los datos reales usando el software Infostat. Las conclusiones y las recomendaciones en el Capítulo 6 y Finalmente, las referencias bibliográficas.

¹Fuente: EFArmania.

Capítulo 2

EL PROBLEMA DE INVESTIGACIÓN

¿Investigación científica libre? El segundo adjetivo es redundante.

Ayn Rand

2.1. Antecedentes de la idea

La participación del sector financiero en el desarrollo económico mundial y nacional es de suma importancia. El sistema financiero es el encargado de captar recursos del público, los cuales posteriormente son colocados a través de créditos a otros agentes económicos. Estos recursos son destinados para inversión en empresas o para cubrir ciertas necesidades de consumo, son estas inversiones parte del motor de crecimiento de la economía. Dicho sector tiene como función principal contribuir con el logro de los objetivos de estabilización y crecimiento económico de un país. Así mismo, para que el sistema financiero pueda cumplir sus objetivos, es necesario que se cum-

plan algunos requisitos a fin de lograr su adecuado funcionamiento, entre los cuales se pueden enumerar los siguientes: Sistema de regulación y supervisión, propiciar un flujo de información adecuado y evitar la información asimétrica y cumplimiento de un marco legal¹. Las entidades financieras se rigen según normas nacionales, supervisadas por la Superintendencia Financiera de Colombia y también según normas internacionales como son los acuerdos de Basilea I, II y III. A pesar del cumplimiento de los parámetros establecidos, en un informe publicado por la CEPAL² en el año 2011, en el financiamiento a las PYMES en América Latina existen problemas basados en la insuficiente información con la que cuentan los bancos para realizar evaluaciones de riesgo, falta de información puntual sobre proyectos de las empresas, ventas reales de las empresas, potencialidad de crecimiento, inexistencia o escasa información de la situación financiera, del estado de resultados integral y de las notas explicativas, y falta de homogeneidad de criterios y políticas para la toma de decisiones en las entidades financieras.

Cómo un ejemplo de modelos diseñados para otorgar crédito a sectores específicos, están los créditos otorgados por FINAGRO, El Fondo para el Financiamiento del Sector Agropecuario. FINAGRO es una entidad que promueve el desarrollo del sector rural colombiano, con instrumentos de financiamiento y desarrollo rural, que estimulan la inversión. Esta entidad es una sociedad de economía mixta del orden nacional, organizada como establecimiento de crédito, con régimen especial, vinculada al Ministerio de Agricultura y Desarrollo Rural y vigilado por la Superintendencia Financiera de Colombia. FINAGRO actúa como entidad de segundo piso, es decir, otorga recursos en condiciones de fomento a las entidades financieras, para que éstas a su vez otorguen créditos a proyectos productivos. Así mismo, para facilitar el acceso al financiamiento,

¹ Tomado de: www.banguat.gob.gt/inveco/notas/articulos/envolver.asp?karchivo=1002&kdisc=si.

² Eliminando Barreras: El financiamiento de las pymes en América Latina. Carlos Ferraro, Evelin Goldstein, Luis Alberto Zuleta y Celso Garrido. Naciones Unidas, 2011.

FINAGRO administra instrumentos para el desarrollo de su proyecto agropecuario³.

Esta entidad tiene una función muy particular y las condiciones de los créditos están establecidas de manera especializada según análisis de cada unidad productiva, y cada crédito nace con un proyecto y es financiado según su destino, difiere en cuanto, plazos, montos, tasas, periodos de amortización, políticas analizadas y definidas con estudios previos de comportamiento de cada unidad productiva.

Describiendo de manera generalizada lo ocurrido con el proceso de otorgamiento de crédito se encontró que históricamente éste ejercicio se hacía por conocimiento de cliente, el criterio más importante estaba representado en la reputación del deudor y el respaldo que tenía en sus propiedades. Posteriormente en los años 70, se inicia el desarrollo de formas más acertadas para otorgar crédito; se hace una segmentación de clientes según características en empresas y personas, sin perder importancia el respaldo patrimonial de los deudores. Luego en la década de los años 90, el sector financiero sufrió una crisis que lo hizo reflexionar dado el nivel de pérdidas no esperadas por mal otorgamiento de créditos. Los indicadores de cartera vencida afectaron al sector, época en la que los bancos deterioraron su imagen por los procesos de remates de propiedades, hecho que obligo a las entidades gubernamentales regular dicho proceso y exigir a las entidades revisar el proceso de otorgamiento de crédito. Fue entonces cuando se incorporaron nuevos criterios de aprobación y se pudo identificar que a un cliente que históricamente se le prestaban sumas importantes amparados en sus activos, con la nueva metodología algunos ya no eran sujetos de crédito. Para el banco lo importante es entonces revisar el otorgamiento de crédito desde el punto de vista de capacidad de pago.

Actualmente sumado al análisis de capacidad de pago se hace necesario el análisis de

³Tomado de: <https://www.finagro.com.co/qui%C3%A9nes-somos/informaci%C3%B3n-institucional>

comportamiento sectorial, se crean políticas dependiendo el segmento, sin embargo no han sido suficientes los esfuerzos y **actualmente en la entidad financiera objeto de éste estudio se ha trasferido la responsabilidad al comercial⁴ de recomendar la operación de crédito. Esta situación evidencia la falta de un criterio uniforme, se encuentran entonces montos de cupos de crédito diferentes para clientes con las mismas características financieras.** Adicionalmente, los niveles de cartera vencida en este segmento están por encima del promedio. También se encuentran casos de algunos clientes que por decisión del comercial no se envían a estudio. Por lo anterior, es necesario tener una herramienta válida que identifique para este tipo de clientes la probabilidad de incumplimiento, para ayudar a tomar decisiones más adecuadas y con un menor riesgo, para lo que **se propone la elaboración del modelo scoring que complemente el análisis actual que realiza la EF Armenia.**

2.2. Situación problema

El proceso de otorgamiento de crédito en una entidad financiera de la ciudad de Armenia al Segmento de Ingenieros Civiles Contratistas (SICC) se ha realizado con los modelos convencionales para PYMES, situación que ha generado un alto nivel de cartera vencida comparado con indicadores de otros segmentos similares de actividades económicas diferentes. En contraste, los resultados de cartera morosa, son muy altos en el segmento estudiado hecho que genera pérdidas para la entidad.⁵

⁴El comercial es el gerente que atiende al cliente y quien hace la gestión para la aprobación del crédito.

⁵Dato de la entidad financiera en estudio a Noviembre, 2016.

2.3. Definición del problema

Se tiene una insuficiencia en el modelo scoring tradicional de la entidad financiera utilizado en el proceso de análisis, evaluación y otorgamiento de créditos para el segmento de Ingenieros Civiles Contratistas SICC.

2.4. Hipótesis o supuestos

Los modelos scoring actuales utilizados en el otorgamiento de crédito al segmento de ingenieros civiles contratistas no son acertados en la medición de probabilidad de incumplimiento.

2.5. Objetivo general

Diseñar un modelo scoring para otorgamiento de crédito comercial a ingenieros civiles contratistas en una entidad financiera del municipio de Armenia.

2.5.1. Objetivos específicos

- Caracterizar el sector de ingenieros civiles contratistas.
- Analizar la reglamentación de entes de control para el cumplimiento de parámetros en la elaboración del modelo scoring.
- Analizar estadísticamente la información de las bases de datos de ingenieros civiles contratistas de una entidad financiera del municipio de Armenia.

- Determinar las variables requeridas para la elaboración del modelo scoring.
- Elaboración del modelo scoring para el segmento de ingenieros civiles contratistas utilizando el método estadístico del Análisis Discriminante.

2.6. Justificación del estudio

Para la entidad financiera objeto de éste estudio, EF Armenia, es sumamente importante analizar al Segmento de Ingenieros Civiles Contratistas de manera diferente. Sus registros de incumplimiento de pagos en fechas acordadas por norma interna del banco son altos comparados con otros sectores similares. Sin embargo, del total de la cartera vencida hay un porcentaje mínimo de cartera castigada⁶. Este hecho permite pensar que existe un problema estructural en el momento del otorgamiento de un crédito, lo que justifica la creación de un modelo scoring diferente, basado en las particularidades de este segmento.

Debido a lo anterior, es de suma importancia estudiar de manera aislada a este tipo de clientes, que por sus ingresos superiores a 400 millones anuales son categorizados y analizados como clientes PYME; pero que por su estructura operativa, de comportamiento y manejo administrativo son similares a los de un profesional independiente. Algunos de ellos carecen de estructura administrativa formal (oficina, asesores, etc.). Sus hábitos de consumo en algunas ocasiones no son acordes al nivel de ingresos cíclico de la actividad. Por este mismo hecho su flujo de caja también es cíclico, ocasionando moras en obligaciones financieras o ante proveedores, debido al mal otorgamiento

⁶CARTERA VENCIDA: Monto total de créditos cuyo capital, cuotas de amortización o intereses no han sido cancelados íntegramente a la entidad en su fecha límite de pago. CARTERA CASTIGADA: Categoría que agrupa a todos los créditos insolutos, después de agotadas las gestiones de cobro ejercidas por todas las instancias del banco que participan en esta gestión. Contablemente significa reconocerlas como pérdida.

to de crédito, donde por ejemplo el plan de pagos no va acorde con su actividad. En correspondencia los ingenieros civiles contratistas se convierten en un segmento que necesita una atención especializada por parte de EFArmania.

El hecho de no optar por construir un modelo particular para clasificar a los clientes de este segmento implicaría para la entidad financiera una pérdida de oportunidad de negocio al no atender clientes buenos o, pérdida económica al otorgar un crédito a un cliente que potencialmente van a incumplir en sus pagos.

2.7. Beneficios que conlleva

El diseño de un modelo scoring para un segmento determinado como el planteado en éste trabajo de investigación permitirá cuantificar el riesgo del SICC con el pronóstico de la probabilidad de incumplimiento del cliente analizado, y posteriormente ayudar a tomar decisiones de aprobación o negación de un crédito, permitiendo a la entidad financiera atender las necesidades de financiación con una mejor predicción de riesgo y consecuentemente disminuir el deterioro de la cartera de préstamos en este sector. Además, la aplicación del modelo permitirá una reducción significativa en los tiempos de ejecución de los distintos procesos de otorgamiento de crédito.

2.8. Limitaciones previsibles

Una de las limitaciones previsibles dentro del estudio es la carencia de información de las solicitudes negadas o no enviadas a análisis para poder evaluarlos dentro del modelo planteado.

2.9. Metodología

Por los objetivos general y específicos planteados, esta investigación se enmarcó dentro del tipo de investigación de carácter descriptivo y de carácter correlacional. De carácter descriptivo porque en este estudio se muestran, narran, reseñan o identifican hechos, situaciones, rasgos de un objeto de estudio, o se diseñan productos, modelos, prototipos, guías, etc. De carácter correlacional porque un aspecto importante de esta investigación consiste en examinar relaciones entre variables o sus resultados, pero en ningún momento explica que una sea causa de la otra[2].

MARCO REFERENCIAL

3.1. Marco Teórico

3.1.1. El Sistema Financiero Colombiano

El sistema económico internacional surgió a partir del año 1945, cuando finalizó la segunda guerra mundial. Se apoya básicamente en tres instituciones: el Fondo Monetario Internacional (FMI), el Banco Internacional de Reconstrucción y Fomento (Banco Mundial) y el Acuerdo de Aranceles y Comercio (GATT).

En junio del año 1944, las Naciones Unidas promueven una conferencia en Bretón Woods, New Hampshire, Estados Unidos. Los representantes de 45 países acuden al llamado y acuerdan crear una organización económica que permitiera la colaboración internacional y pudiera evitar que se repitiera las equivocadas medidas de política económica que conllevaron a la gran depresión de los años 30. El organismo creado, El Fondo Monetario Internacional (FMI), tiene como objetivo principal evitar las crisis

en los sistemas monetarios, alentando a los países a adoptar medidas de política económica bien fundadas. Como su nombre indica, la institución es también un fondo al que los países miembros que necesiten financiamiento temporal pueden recurrir para superar los problemas de balanza de pagos. Otro objetivo es promover la cooperación internacional en temas monetarios internacionales y facilitar el movimiento del comercio a través de la capacidad productiva (Finzansun406, 2015).

Historia del sector financiero colombiano



Figura 3.1: Billeto de un peso colombiano.

La historia del sector financiero en un país está siempre ligada a la de su sociedad y su economía. Por ello, la historia del sector financiero colombiano, responde a la evolución de las ideas y de la economía colombiana a lo largo de doscientos años de vida independiente. Se inicia con los rudimentarios esquemas de crédito heredados de la colonia¹ y termina con el refinado sistema actual, después de vivir las experiencias de la banca libre, la centrada en la operación de corto plazo, la de fomento, la estatal y la especializada, hasta llegar a la banca universal de hoy en día.

¹La época de la colonia: Los siglos que transcurrieron entre 1550 y 1810 se han denominado *época colonial* debido a la presencia y al dominio político por parte de los españoles en lo que actualmente comprende el territorio de Colombia. Durante este tiempo se formó en América una sociedad en la que las costumbres, la lengua y la religión traídas por los españoles se mezclaron con la cultura indígena y, más tarde, con la africana. Así se conformó lo que hoy en día es Hispanoamérica. Tomado de: <http://lacoloniaeci2006.blogspot.com.co/>

En la colonia hubo mecanismos de ahorro y crédito. Existió, quien recibía dinero a interés y lo prestaba para ganarse el margen de intermediación. La Iglesia Católica, confundida con la corona española, estableció maneras para otorgar crédito a través de una especie de hipoteca de los inmuebles: la “capellanía”, mediante la cual algunos bienes quedaban sujetos al cumplimiento de misas y de otros compromisos religiosos.

La idea de establecer un banco estatal, para financiar al gobierno, se plasmó por primera vez en el Congreso Constituyente de Cúcuta en 1821, pero no prosperó. Al nacer la República de Colombia, el crédito público dependió de los préstamos de los comerciantes antioqueños al gobierno central, de las contribuciones forzosas y de los empréstitos ingleses. En 1833, una ley consideró que la actividad bancaria formaba parte de la comercial; en 1839, Judas Tadeo Landinez² constituyó en Bogotá un establecimiento mercantil y financiero que un par de años más tarde se convirtió en una compañía de “giro y descuento”. Un banco privado que, al sobre extenderse en la compra de vales del gobierno y en distintas inversiones, dio lugar a la primera bancarrota en la historia del país.

La banca comercial propiamente dicha se iniciaría en los años sesenta del siglo XIX cuando la economía comenzó a vincularse con el extranjero a través de las exportaciones agrícolas. Se implantó el modelo conocido como de “banca libre” bajo el cual, en ausencia de un banco central y de la regulación estatal, los mismos bancos eran responsables de que se mantuviera la confianza del público en los billetes convertibles en oro, que emitían en un ambiente de libre competencia. En 1870 se creó el Banco de

²Judas Tadeo Landinez, nació en el municipio de Samacá, estudió en Tunja y luego en Bogotá donde se recibió de abogado. Fue diputado en épocas del General Santander, representante al Congreso, funcionario de hacienda en Tunja, rector del Colegio Boyacá de Tunja, Ministro de Hacienda, Ministro del Interior y Ministro de Relaciones Exteriores entre otros cargos que ejerció especialmente en el Gobierno del presidente Márquez. Su amplia trayectoria política le sirvió de base y respaldo para incursionar en el mundo financiero, en el que se dice inicio sus negocios con un capital de \$22.000. Tomado de: <http://ruborpuebloscundinamarca.blogspot.com.co/2012/05/judas-tadeo-landinez-y-su-imperio.html>.

Bogotá y en 1874 el Banco de Colombia hoy Bancolombia.

El período de la “banca libre” pura duró poco tiempo. En 1880, la ley autorizó al presidente Rafael Núñez³ para organizar el Banco Nacional como agente financiero del gobierno central; los bancos privados fueron obligados en ese momento a aceptar los billetes emitidos por el Banco Nacional. Este nunca fue un banco central; fue un banco del gobierno, que prestaba dinero y que comenzó a emitir billetes sin respaldo en oro lo que habría de provocar un escándalo político de grandes dimensiones. En los años noventa del siglo XIX el gobierno decidió cerrarlo. Vino después la guerra de los mil días⁴ y un gran desorden monetario que desembocó en la hiperinflación de principios del siglo XX.

Solamente hasta 1923 el gobierno del presidente Pedro Nel Ospina⁵ invitó a una misión extranjera, presidida por el profesor Edwin Kemmerer⁶ de la Universidad de Princeton en los Estados Unidos para que organizara la moneda y las finanzas públicas nacionales; se creó en ese año el Banco de la República. En 1923 se creó, también, la Superintendencia Bancaria para regular y supervisar el incipiente sistema bancario nacional, que ya se había visto apuntalado por el establecimiento en Colombia de sucursales de bancos extranjeros, como el First National Citibank, el Royal Bank of Canadá o el Banco

³Cartagena de Indias, Colombia, 1825 - 1894. Político colombiano, presidente de Colombia entre 1880-1882, 1884-1886 y 1886-1888. Tomado de: <http://www.biografiasyvidas.com/biografia/n/nunez-rafael.htm>.

⁴La guerra de los mil días, fue una guerra civil de Colombia disputada entre el 17 de octubre de 1899 y el 21 de noviembre de 1902, en un principio entre el Partido Liberal y el gobierno del Partido Nacional en cabeza del presidente Manuel Antonio Sanclemente, quien fue derrocado el 31 de julio de 1900 por José Manuel Marroquín Ricaurte, representante del Partido Conservador, en alianza con el liberal Aquileo Parra. Tomado de: <http://www.colombia.com/colombia-info/historia-de-colombia/epoca-contemporanea/despertar-siglo-xx/la-guerra-de-los-mil-dias/>

⁵Bogotá, 1858 - Medellín, 1927. Militar y político colombiano que fue presidente de Colombia entre 1922 y 1926. Tomado de: <http://www.biografiasyvidas.com/biografia/o/ospina-pedro.htm>.

⁶Político y economista estadounidense. Nació en Scranton, Pensilvania el 29 de julio de 1875 y murió en Princeton, Nueva Jersey el 16 de diciembre de 1945. Era considerado uno de los mejores economistas de su época, lo que hizo que muchos países, especialmente americanos, solicitaran sus servicios como asesor en temas de economía y finanzas. Tomado y adaptado de: <http://www.mcnbiografias.com/app-bio/do/show?key=kemmerer-edwin-walter>

Francés e Italiano.

El banco central de emisión era una necesidad para la expansión de la economía. A partir de 1923 comenzaron a entrar los 25 millones de dólares provenientes de la indemnización que el gobierno de los Estados Unidos reconoció a Colombia por la pérdida del canal de Panamá, que había ocurrido en 1903, lo cual abrió el crédito externo para el país. La producción y las exportaciones de café se incrementaban aceleradamente y había comenzado un proceso de industrialización y de urbanización. En 1928 se crearía la Bolsa de Bogotá para apoyar la capitalización de las empresas manufactureras y los mismos bancos. Los bancos comerciales se concentraron hasta los años cincuenta en el crédito de corto plazo para agricultores e industriales. En los principios de los años treinta, se crearon dos entidades bancarias que habrían de convertirse en emblemas colombianos en el siglo XX: La Caja de Crédito Agrario Industrial y Minero y el Banco Central Hipotecario. Ambas se liquidaron setenta años más tarde, como consecuencia de la crisis financiera de fin de siglo.

El crédito de largo plazo, para inversión en la industria y en la agricultura, no sería realidad sino hasta 1951 cuando una reforma financiera permitió que los bancos comerciales obtuvieran recursos para ese fin en el Banco de la República. Se inició en ese año la era del “crédito de fomento”, que se prolongaría hasta 1991.

En la primera mitad de los años cincuenta se fundaron los bancos estatales orientados a la financiación de sectores específicos, como el Banco Popular, el Banco Ganadero y el Banco Cafetero. Y a finales de esa década y principios de los sesenta hicieron su aparición las corporaciones financieras con el objeto de realizar inversiones en los sectores productivos.

En la década de los años setenta se autorizó la creación de otra entidad especializa-

da, esta vez en la captación de recursos de ahorro para destinarlos a la construcción de vivienda: las corporaciones de ahorro y vivienda, eje del sistema de ahorro de valor constante, que se conoció hasta fines del siglo como el sistema UPAC⁷. Los nombres de estas entidades se popularizaron rápidamente, como Davivienda, Granahorrar, Concasa, Conavi, Corpavi, Colpatria, Colmena o Ahorramás. A finales de los setenta, por último, se formalizaron los intermediarios financieros extra bancarios, dedicados a la financiación de crédito de consumo, al transformarse en compañías de financiamiento comercial.

El sistema financiero se fue conformando a partir de los años cincuenta del siglo XX como uno de entidades especializadas en la financiación de los distintos sectores. En los años noventa, sin embargo, se inició la transición hacia un sistema de banca universal, centrado en los bancos comerciales. Las crisis financieras (de los ochenta y finales del siglo, la más profunda de todas) dieron lugar a procesos de cierre de entidades y de fusiones y adquisiciones entre ellas, lo cual facilitó la consolidación y el reforzamiento del sistema en los principios del siglo XXI. La apertura a la inversión extranjera en el sector financiero de los años noventa impulsó, igualmente, la transformación del sector financiero en Colombia.

En la actualidad el sistema financiero colombiano cuenta con un menor número de entidades, las entidades especializadas prácticamente han desaparecido y la regulación, supervisión y vigilancia de las existentes permiten que se mantenga la confianza del público en esas entidades, a pesar de la frecuente crítica de la cual son objeto, en todo el mundo, los bancos comerciales. Se transitó por un largo camino, lleno de difi-

⁷UPAC son las iniciales de unidad de poder adquisitivo constante. Este sistema tuvo sus orígenes en Brasil, y fue implantado en Colombia en el año 1972 durante la presidencia de Misael Pastrana Borrero. La UPAC tenía como principales objetivos los de mantener el poder adquisitivo de la moneda y ofrecer una solución a los colombianos que necesitaran tomar un crédito hipotecario de largo plazo para comprar vivienda. Tomado de: http://www.banrepcultural.org/blaavirtual/ayudadetareas/economia/el_upac_y_la_uvr.

cultades y obstáculos, hasta llegar a la sólida y eficiente organización financiera de hoy en día⁸.

3.1.2. Modelos scoring

Esta sección hace una presentación general de lo que se entiende por modelación scoring para control de riesgo. Se entrega una visión teórica general sobre las técnicas estadísticas para la creación de los modelos de credit scoring, especialmente el análisis discriminante dejando las bases para la modelación y aplicación de ésta herramienta sobre la base de datos suministrada por EFArmenia. Éste estudio empírico, tiene como objetivo analizar y valorar la morosidad, como forma de manifestación de dicho riesgo.

El objetivo de un modelo de credit scoring se centra en la medición y/o pronóstico del riesgo de crédito. Las técnicas y modelos utilizados por las entidades financieras en su gestión han seguido una evolución acorde con la exigencia, cada vez mayor, de los gestores empresariales demandantes de herramientas para la toma de decisiones en la asunción de riesgos. Aunque en sus inicios los modelos de valoración del riesgo de crédito eran de corte univariante, contemplando el riesgo crediticio en función de una sola variable, o de considerar varias, analizando cada una por separado, el hecho de que la asunción de riesgos dependa de múltiples factores, provocó que las técnicas evolucionaran hacia modelos de tipo multivariante. Dentro de éstos, se distinguen en primer lugar, aquéllos de carácter subjetivo, basados en el criterio y juicio de analistas expertos, quienes utilizan sus conocimientos y experiencias acumuladas a través de años de trabajo para evaluar el riesgo de crédito y tomar decisiones al respecto. **Con el**

⁸Esta sección 3.2 fué tomada y adaptada de: “Un rápido recorrido por la historia del sector financiero en Colombia” de Carlos Caballero Argáez, revista DINERO 2010.

fin de superar los inconvenientes que tales modelos presentan, en cuanto a la subjetividad de las decisiones, se desarrollaron después modelos objetivos de evaluación, basados en métodos estadísticos, que consiguen que distintos analistas, a partir de idénticos elementos o factores, lleguen a las mismas decisiones.

Los métodos estadísticos que las entidades financieras aplican para el análisis de sus clientes con el propósito de minimizar el riesgo y maximizar la rentabilidad en la colocación de créditos son principalmente el **análisis discriminante** y la **regresión logística** como herramienta de la metodología llamada *credit scoring*. Generalmente las herramientas cuantitativas de medición del riesgo de crédito descansan sobre esta metodología acompañada también de otra conocida como *rating*⁹. Estas metodologías permiten calificar el riesgo mediante una puntuación o una clasificación en grupos de riesgo.

Los modelos de *credit scoring*, también llamados *score-cards* o *classifiers*, son métodos algorítmicos que de manera automática evalúan el riesgo de crédito para un solicitante de financiamiento o de alguien que ya es cliente de la entidad. Su razón principal es medir el riesgo de impago en forma individual, ya que se enfocan en el riesgo de incumplimiento del individuo o empresa, independientemente de lo que ocurra con el resto de la cartera de préstamos. Este es uno de los aspectos en los que se diferencian de otras herramientas de medición del riesgo de crédito, como son los modelos de cartera y los VaR marginales, que tienen en cuenta la correlación de la calidad crediticia de los deudores de una cartera de préstamos.

⁹El rating, o calificación crediticia, es un indicador de solvencia, elaborado por una agencia especializada, cuyo fin es el de informar, a terceros interesados, de la calidad crediticia o solvencia de una sociedad. Un rating es una calificación de la capacidad solvente de una compañía. De forma resumida, juzga la capacidad o probabilidad, que tiene una empresa de atender, en la forma y el tiempo correcto, a los intereses y el principal del conjunto de sus deudas bancarias y comerciales. Tomado de: <http://www.antoniograndiodopico.es/archivos/Varios/RATING.pdf>.

En una primera aproximación a los mismos, se los puede definir como “métodos estadísticos utilizados para clasificar a los solicitantes de crédito, o incluso a quienes ya son clientes de la entidad evaluadora, entre las clases de riesgo *bueno* y *malo*” (Hand y Henley (1997)). Aunque originalmente en los años 70 se basaban en técnicas estadísticas (en particular, el análisis discriminante), en la actualidad también están basados en técnicas matemáticas, econométricas y de inteligencia artificial. En cualquier caso, los modelos de credit scoring emplean principalmente la información del evaluado contenida en las solicitudes de crédito y/o en fuentes internas y/o externas de información.

El resultado de la evaluación se refleja en la asignación de alguna medida que permita comparar y ordenar a los evaluados en función de su riesgo, a la vez que cuantificarlo. Por lo general, los modelos de credit scoring le asignan al evaluado un puntaje o score, o una calificación, clasificación o rating. Algunos métodos los asignan a grupos, en donde cada grupo tiene un perfil de riesgo distinto; sin embargo, en la práctica esto equivale a una calificación. A su vez, estos ordenamientos de los deudores permiten obtener estimaciones más concretas del riesgo; en general se busca obtener alguna estimación de la probabilidad de incumplimiento del deudor (PD, por probabilidad de default) asociada a su score, rating o calificación. Esta estimación puede obtener directamente el score en el caso de los modelos econométricos, o también en función de la tasa de incumplimiento (TD, por tasa de default) histórica observada en el grupo de deudores con la misma calificación o score similar.

Métodos utilizados

Para evaluar el riesgo de crédito, se utiliza una gran variedad de herramientas estadísticas como son: Análisis discriminante, Regresión lineal, Regresión logística o Logit,

Modelos Probit, Métodos no paramétricos de suavizado, Métodos de programación matemática, Modelos basados en cadenas de Markov, Algoritmos de particionamiento recursivo (árboles de decisión), Sistemas expertos, Algoritmos genéticos, Redes neuronales y, el juicio humano (la decisión de un analista acerca de otorgar un crédito), entre otros.

Entre todas las metodologías disponibles, los modelos probit, junto con la regresión lineal y logística, el análisis discriminante y los árboles de decisión, se encuentran entre los métodos más usados en la industria para confeccionar estos modelos (Gutiérrez Girault, 2007).

3.1.3. Modalidades de contratación pública

Los contratistas son personas naturales o jurídicas, que proveen bienes y servicios a entidades públicas de forma independiente o de forma plural en consorcio y/o unión temporal mediante procesos de contratación pública y bajo la reglamentación del Estatuto General de Contratación de la Administración Pública (Ley 80 de 1993 y la Ley 1150 de 2007). De acuerdo a los servicios que prestan a las diversas entidades públicas la entidad financiera los cataloga según el perfil de cada uno en:

- Perfil 1: Dedicado a desarrollar proyectos de obra por medio de la participación de contratos públicos. Estos trabajos pueden representar la totalidad de la obra, o bien sea parte de ella. Para ello, es común que el contratista se apoye en otra persona u organizaciones para que realicen determinado tipo de trabajos especializados; a ellos se les llama Subcontratistas.
- Perfil 2: Provee bienes y servicios como: dotación, uniformes, alimentación, combustibles, salud, interventoría, entre otros.

- Perfil 3: Profesional independiente que presta servicios de consultoría y asesoría por medio de contratos públicos o privados.

Estos perfiles de contratistas definidos por la entidad financiera ejecutan diferentes contratos, clasificados por la administración pública en 7 tipos de contratos estatales descritos a continuación: Contrato de consultoría, Contrato de prestación de servicios, Encargos fiduciarios y Fiducia pública, Contrato de obra, Contrato de concesión, Contrato marco de precios, Asociación público privada. A continuación se presentan con más detalle.

1. **Contrato de consultoría:** Son contratos que celebran las entidades estatales referidas a los estudios necesarios para la ejecución de proyectos de inversión, estudios de diagnóstico, pre-factibilidad o factibilidad para programas o proyectos específicos, diseños e interventoras, así como asesorías técnicas de coordinación, control y supervisión.
2. **Contrato de prestación de servicios:** Son contratos que celebran las entidades estatales para desarrollar actividades relacionadas con la administración o funcionamiento de la entidad. Estos contratos solo podrán celebrarse con personas naturales cuando dichas actividades no puedan desarrollarse con personal de planta o requieran conocimientos especializados.
3. **Encargos fiduciarios y fiducia pública:** Los encargos fiduciarios que celebran las entidades estatales con las sociedades fiduciarias autorizadas por la Superintendencia Financiera tienen por objeto la administración o el manejo de los recursos vinculados a los contratos que tales entidades celebren.
4. **Contrato de obra:** Son contratos que celebran las entidades estatales para la construcción, mantenimiento, instalación y, en general, para la realización de cual-

quier otro trabajo material sobre bienes inmuebles, cualquiera que sea la modalidad de ejecución y pago. Estos podrán contar con recursos públicos de hasta el 20% del presupuesto estimado de inversión inicial del proyecto. En este caso la selección del inversionista será realizada a través de licitación pública. Cuando la financiación de la totalidad del proyecto esté a cargo de los privados opera el mecanismo de selección abreviada.

5. **Contrato de concesión:** Son contratos que celebran las entidades estatales con el objeto de otorgar a una persona llamada concesionario, la prestación, operación, explotación, organización o gestión de un servicio público; o la construcción, explotación o conservación total o parcial de una obra o bien destinados al servicio o uso público, así como aquellas actividades necesarias para la adecuada prestación o funcionamiento de la obra o servicio por cuenta y riesgo del concesionario y bajo la vigilancia y control de la entidad contratante, a cambio de una remuneración que puede consistir en derechos, tarifas, tasas, valoración, o en la participación que se le otorgue en la explotación del bien, o en una suma periódica, única o porcentual y en general, en cualquier otra modalidad de contraprestación que las partes acuerden.
6. **Contrato marco de precios:** Los acuerdos o convenios marco de precios son una herramienta para que el Estado agregue demanda y centralice decisiones de adquisición de bienes, obras o servicios para: producir economías de escala, incrementar el poder de negociación del Estado y compartir costos y conocimientos entre las diferentes agencias o departamentos del Estado. El Acuerdo Marco de Precios es un contrato entre un representante de los compradores y uno o varios proveedores, que contiene la identificación del bien o servicio, el precio máximo de adquisición, las garantías mínimas y el plazo mínimo de entrega, así como las condiciones a través de las cuales un comprador puede vincularse al acuer-

do. Generalmente, los compradores se vinculan a un acuerdo marco de precios mediante una manifestación de su compromiso de cumplir las condiciones del mismo y la colocación de una orden de compra para la adquisición de los bienes o servicios previstos en el acuerdo.

7. **Asociaciones Público Privadas APP:** Son un mecanismo que permite vincular al sector privado para proveer bienes y servicios públicos asociados a una infraestructura. La Asociación Publico Privada involucra, dentro del contrato de largo plazo, la retención y transferencia de riesgos entre las partes y la definición de los mecanismos de pago, relacionados con la disponibilidad y el nivel de servicio del bien provisto. Hay dos tipos de APP: iniciativa pública e iniciativa privada:

- APP Iniciativa Pública: La idea conceptual del proyecto es estructurado por la entidad pública con participación del sector privado. La fuente de pago del proyecto puede ser a través de aportes de recursos públicos, de la explotación económica de la APP o una combinación de estas.
- APP de Iniciativa Privada: La idea conceptual y propuesta proviene del sector privado. En este caso el privado tiene la responsabilidad de realizar, por su propia cuenta y riesgo, la estructuración del proyecto sin obligación del público a reconocerle los costos asociados.

Existen diferentes modalidades de selección para ser contratista del estado, a continuación se relacionan con sus principales generalidades.

1. **Licitación pública:** Es un proceso por medio del cual las entidades estatales abren una convocatoria para suplir una necesidad específica. La entidad contratante establece previamente las características del bien o servicio que está solicitando, a las cuales se deben ajustar quienes estén interesados en participar en la licita-

ción, se podría resumir en el siguiente proceso: la entidad contratante realiza los estudios específicos y la estructuración de los proyectos. publica pre-pliegos de condiciones y programa visitas a las obras, ejecuta reuniones informativas y recibe observaciones a los pre-pliegos, realiza modificaciones a los pre-pliegos y publica pliegos definitivos, recibe las ofertas técnicas y económicas de los proponentes, ejecuta audiencia para adjudicar el contrato y firma el contrato con el proponente que resulte ganador en la licitación.

2. **Selección abreviada:** Es una modalidad de selección objetiva prevista para aquellos casos en que por las características del objeto a contratar, las circunstancias de la contratación o la cuantía o destinación del bien, obra o servicio, pueden adelantarse procesos simplificados para garantizar la eficiencia de la gestión contractual. Esta modalidad de selección se usa para la prestación de servicio de salud y los bienes y servicios para la defensa y la seguridad nacional, entre otros.
3. **Selección abreviada para la adquisición de bienes uniformes (Subasta):** La Ley 1150 de 2007 estableció el mecanismo de subasta inversa para la adquisición o suministro de bienes y servicios de características técnicas uniformes y de común utilización por parte de las entidades, que corresponden a aquellos que poseen las mismas especificaciones técnicas, con independencia de su diseño o de sus características descriptivas, y comparten patrones de desempeño y calidad objetivamente definidos. La subasta puede adelantarse de manera electrónica o presencial y en cualquiera de los casos deben ser publicados los documentos en el Secop¹⁰.

¹⁰El Secop es un sistema electrónico que entre otras funcionalidades, permite a las entidades estatales cumplir con las obligaciones de publicidad de los diferentes actos expedidos en los procesos contractuales y permite a los interesados en participar en los procesos de contratación, proponentes, veedurías y a la ciudadanía en general, consultar el estado de los mismos. Es importante resaltar que según el documento Conpes 3620 de 2009, el Secop se creó como un incentivo para el uso del comercio electrónico, cuyo desarrollo se realizará a futuro una vez se de inicio a la Fase transaccional la cual permitirá “llevar a cabo la gestión pública contractual a través de medios electrónicos”. Tomado de: http://www.aplicaciones-mcit.gov.co/secop/que_es_secop.html.

4. **Selección abreviada del literal h del numeral 2 del artículo 2 de la Ley 1150 de 2007:** Hace referencia a los contratos de las entidades, a cuyo cargo se encuentre la ejecución de los programas de protección de personas amenazadas, programas de desmovilización y reincorporación a la vida civil de personas y grupos al margen de la ley, incluida la atención de los respectivos grupos familiares, programas de atención a población desplazada por la violencia, etc.
5. **Selección abreviada servicios de salud:** Sin perjuicio de lo dispuesto en la Ley 100 de 1993 y en la Ley 1122 de 2007, la Ley 1150 de 2007 definió como una de las causales de selección abreviada, la celebración de contratos para la prestación de servicios de salud.
6. **Concurso de méritos:** Modalidad prevista para la selección de consultores o proyectos, en la que se podrán utilizar sistemas de concurso abierto o de precalificación. En este último caso, la conformación de la lista de precalificados se hará mediante convocatoria pública, permitiéndose establecer listas limitadas de oferentes utilizando criterios de experiencia, capacidad intelectual y de organización de los proponentes, según sea el caso. Bajo esta modalidad el precio de la propuesta no constituye factor de escogencia.
7. **Concurso de méritos abierto:** El concurso de méritos por el sistema de concurso abierto se desarrolla prescindiendo de los procedimientos de precalificación, es decir que podrán participar todos los interesados para ésta modalidad de selección y de acuerdo con los documentos que deben ser publicados.
8. **Concurso de méritos con lista corta:** Consiste en la conformación de una lista limitada de oferentes para un solo proceso de concurso de méritos. La entidad conformará la lista corta con un número plural de precalificados que no podrá exceder de seis cuando se deba presentar una propuesta técnica detallada, ni de diez cuando se deba presentar una propuesta técnica simplificada.

9. **Lista multiusos:** Consiste en la conformación de una lista limitada de oferentes (precalificación) para varios concursos de méritos determinados o determinables. En esta solo se publican los documentos del proceso que se adelante para la conformación de la lista.
10. **Concurso de méritos con lista multiusos:** Una vez conformada la lista multiusos, en esta opción se publican los datos del proceso de concurso.
11. **Contratación mínima cuantía:** Corresponde a la publicación de procesos cuya adquisición de bienes, servicios y obras cuyo valor no exceda del 10% de la menor cuantía de la entidad contratante, independientemente de su objeto.
12. **Contratación directa:** Es el proceso excepcional por medio del cual el estado no contrata a través de la licitación o el concurso sino que teniendo en cuenta la naturaleza o la cuantía del objeto a contratar, lo hace directamente. Cuando se contrata a través de este medio, se hace por situaciones especiales que hacen inviable el proceso licitatorio.

3.2. Marco Conceptual

3.2.1. Riesgo de crédito

Dentro de las particularidades del cliente de una entidad financiera, especialmente la referente a su actividad económica, existen dos factores inherentes a la misma como son la incertidumbre y la aleatoriedad. Estas dos situaciones son las que se constituyen en la esencia del riesgo crediticio, evaluado por la entidad financiera para ayudar a la toma de decisiones con respecto a la otorgación o no de un crédito financiero. Este riesgo para la entidad financiera se traduce en la contingencia, probabilidad o proxi-

midad de un daño o peligro, en concreto, de sufrir una pérdida.

En el caso de las entidades financieras, el manejo del riesgo va acompañado de la necesidad de rentabilidad. Por lo tanto estos dos aspectos van de la mano en la actividad de otorgamiento de crédito.

El riesgo de crédito es el más importante dentro del quehacer de una entidad financiera, pues éste tipo de riesgo es el que está inmerso en los procesos de negociación con sus clientes. El riesgo de crédito podría definirse, de forma genérica, como el *posible quebranto que originaría a la entidad una variación en las condiciones y características de una contrapartida, que pudiese alterar la capacidad de ésta para cumplir con los términos contractuales de una operación*¹¹. El riesgo puede definirse, de forma más completa, como la exposición a la posibilidad de pérdidas económicas y financieras, de daños en las cosas y en las personas o de otros perjuicios como consecuencia de la incertidumbre que origina el llevar a cabo una determinada acción. Esta incertidumbre es una de las características fundamentales del riesgo, se presenta en tres aspectos (Pecunia, 1 (2005), pp. 175-199):

- En cuanto a la posibilidad de que se materialice una pérdida efectiva.
- Sobre la cuantía de la pérdida.
- Sobre el instante en que puede materializarse dicha pérdida.

El riesgo de crédito es uno de los muchos a los que se enfrenta una entidad financiera. Entre ellos está el riesgo de mercado, riesgo operacional, riesgo de liquidez y riesgo legal. A continuación se presenta con más detalle el riesgo de crédito que es aquel en

¹¹Veáse “Aplicación del Análisis Discriminante y Regresión Logística en el estudio de la morosidad en las entidades financieras. Comparación de resultados” de M. Jesús Mures Quintana, Ana García Gallego y M. Eva Vallejo Pascual. Pecunia, 1 (2005), pp. 175-199.

el que se centrará este trabajo de investigación.

En la relación con los clientes del segmento de ingenieros civiles contratistas (SICC), el riesgo que acompaña a toda operación que las entidades formalizan con ellos, con el fin de concederles crédito, se puede determinar gracias a tres etapas:

- **Etapa 1**-En el estudio de la operación: la entidad financiera tiene que realizar un análisis riguroso del cliente antes de tomar la decisión, por lo que dedica un cierto tiempo a la evaluación de la operación, a partir de la documentación e información necesaria.
- **Etapa 2**-En la formalización de la operación: se debe actuar de forma rigurosa, estableciendo los controles de seguridad adecuados.
- **Etapa 3**-Entre la formalización y la cancelación: durante la vida de la operación, la entidad financiera debe realizar el oportuno seguimiento de la misma y del cliente.

En el marco de estas etapas se encuentran factores característicos de la operación, como son por ejemplo: el plazo o vencimiento de la operación, las garantías aportadas por el cliente¹², estudio de los históricos crediticios del cliente (interna y externa) como son las centrales de riesgo¹³.

¹²Garantías exigidas por la entidad financiera con ánimo de reducir el riesgo de pérdida en caso de no pago.

¹³En Colombia dos instituciones manejan la información crediticia de las personas y funcionan como una base de datos para el sistema financiero: Datacrédito y Cifin. Tomado de <http://www.finanzaspersonales.com.co/credito/articulo/que-centrales-riesgo-para-que-sirven/51380>.

3.2.2. Introducción al Análisis Discriminante

El Análisis Discriminante es una técnica estadística multivariante cuya finalidad es analizar si existen diferencias significativas entre grupos de objetos respecto a un conjunto de variables medidas sobre los mismos, y en el caso de existir diferencias explicar en qué sentido se dan y proporcionar procedimientos de clasificación sistemática para las observaciones en cada uno de los grupos considerados. La técnica tiene como objetivos principales los siguientes:

1. **Análisis discriminante descriptivo:** Describe las diferencias existentes entre grupos, si es que las hay, con base en los valores que toman ciertas variables para cada una de las unidades que integran los distintos grupos considerados.
2. **Análisis discriminante predictivo:** Clasifica las observaciones de acuerdo a sus valores y a las características determinantes en cada uno de los grupos conformados, garantizando una completa homogeneidad dentro del grupo y heterogeneidad entre grupos. La predicción de pertenencia a los grupos se basa en el uso de una o más ecuaciones matemáticas, llamadas funciones discriminantes, que permitan la clasificación de nuevos casos a partir de la información que poseemos sobre ellos a determinados grupos diferenciados y conocidos.

Entre los más apreciable de esta técnica está la claridad en la interpretación que puede hacerse de los resultados, en el sentido de que un número amplio de variables puede ser reducido a un pequeño número de dimensiones capaces de explicar las diferencias fundamentales observadas entre los grupos¹⁴.

Para efectuar el análisis discriminante es necesario considerar los siguientes supues-

¹⁴Tomado de: http://ruc.udc.es/dspace/bitstream/handle/2183/16382/DeDiegoLago_SantiagoInigo_TFG_2015.pdf?sequence=2

tos:

1. Se tiene una variable categórica dependiente a discriminar y el resto de variables son cuantitativas independientes llamadas variables discriminantes.
2. Se requieren al menos dos grupos y para cada grupo dos o más casos.
3. El número de variables discriminantes debe ser menor que el número de objetos o casos menos 2, es decir, si se tiene X_1, X_2, \dots, X_p variables discriminantes, entonces $p < n - 2$, donde n es el número de objetos.
4. Ninguna variable discriminante puede ser combinación lineal de otras variables discriminantes.
5. El número máximo de funciones discriminantes es el mínimo entre el número de variables y el número de grupos menos 1, es decir, con q grupos se tiene $q - 1$ funciones discriminantes.
6. Las matrices de covarianzas dentro de cada grupo deben ser aproximadamente iguales.
7. Las variables continuas deben seguir una distribución normal multivariante.

Técnicamente, se puede decir que el análisis discriminante trata de encontrar funciones cuyos valores separen o discriminen lo más posible a los grupos existentes. Estas funciones, denominadas funciones o ejes discriminantes, serán combinaciones lineales de las variables originales.

Modelo matemático

Partiendo de q grupos donde se asignan a una serie de objetos y de p variables medidas sobre ellos (X_1, X_2, \dots, X_p) , se trata de obtener para cada objeto una serie de puntuaciones que indican el grupo al que pertenecen:

$$y_1, y_2, \dots, y_m$$

con $m = \min [q - 1, p]$, de modo que sean funciones lineales de las X_i , con $i = 1, \dots, p$ así:

$$\begin{cases} Y_1 &= \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \dots + \beta_{1p}X_p + e_1 \\ Y_2 &= \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \dots + \beta_{2p}X_p + e_2 \\ &\vdots \\ Y_m &= \beta_{m0} + \beta_{m1}X_1 + \beta_{m2}X_2 + \dots + \beta_{mp}X_p + e_m, \end{cases}$$

tales que discriminan o separan lo máximo posible a los q grupos y donde e es el término conocido como perturbación y en él se conglomeran todas aquellas variables no tenidas en cuenta en el modelo y, además aquellos errores experimentales no controlables.

Para el caso que interesa, Y es la variable que expone el pago o no pago (default) y las X_i , con $i = 1, \dots, p$ son las distintas variables que se suponen relevantes para explicar Y .

El vector

$$B = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp}),$$

denota una lista de parámetros que recogen la magnitud con que las variaciones en los valores de las variables X se transmiten a variaciones en la variable Y . Estos coeficientes B , se eligen de tal forma que se consiga la máxima heterogeneidad entre grupos, es

decir, tratando de que los valores que toman las funciones discriminantes Y_j en cada uno de sus grupos sean lo más diferentes posibles.

Estadísticamente, este criterio equivale a maximizar la varianza “entre grupos” frente a la varianza “dentro de grupos”. Por lo tanto, los coeficientes $\beta_0, \beta_1, \dots, \beta_k$ se eligen de tal forma que se consiga maximizar el valor del cociente:

$$\pi = \frac{\text{Varianza entre grupos}}{\text{Varianza dentro de grupos}}.$$

Si la varianza “entre grupos” es grande, es decir si hay grandes diferencias entre los valores que toma la función Y en los distintos grupos, pero la varianza “dentro de grupos” es pequeña, es decir, los valores de Y para variables de un mismo grupo son muy similares, entonces se dice que la función discriminante separa bien a los grupos, que serán, internamente muy homogéneos y a la vez muy diferentes entre sí.

Hay que señalar que el número de funciones que pueden obtenerse es el mínimo entre el número de variables explicativas disponibles y el número de grupos menos uno. Estas funciones se obtienen de forma sucesiva en función de su capacidad discriminatoria. Así, la primera función discriminante, será de la forma:

$$Y_1 = \beta_{10} + \beta_{11}X_1 + \beta_{12}X_2 + \dots + \beta_{1p}X_p + e_1,$$

que será la que tenga mayor poder discriminatorio, es decir, la que mejor separe los grupos. La segunda función, que vendrá definida por:

$$Y_2 = \beta_{20} + \beta_{21}X_1 + \beta_{22}X_2 + \dots + \beta_{2p}X_k + e_2,$$

será la siguiente en capacidad discriminatoria, además, estará incorrelacionada con la función anterior Y_1 . A los valores de estas funciones para cada uno de los individuos de la población se les denomina *puntuaciones discriminantes*.

En ciertas ocasiones, la capacidad discriminadora de la primera función Y_1 , es tan grande que la información añadida por la segunda función Y_2 apenas es relevante y esta se ignora, ya que su contribución a la separación entre los grupos no es significativa.

Una vez obtenidas las funciones discriminantes, el objetivo es establecer la contribución relativa de las distintas variables a la discriminación, o lo que es lo mismo, determinar cuáles son las variables que más contribuyen a discriminar entre un grupo y otro.

NOTA IMPORTANTE: *Dado que en éste trabajo se diferencia entre dos grupos de clientes, es decir, los solvente y los insolventes o morosos, se aplica el análisis discriminante simple, en el que la variable dependiente Y es dicotómica, siendo sus categorías cada una de los grupos definidos a priori y, por lo tanto, se obtiene una única función discriminante.*

En cuanto a las variables explicativas o predictoras, éste método impone la condición de que sean cuantitativas o métricas, **aunque pueden introducirse también variables cualitativas o categóricas, sin mas que transformarlas previamente en tantas variables ficticias o *dummy*** como modalidades tenga la variable original menos una. De ésta forma, se facilita el tratamiento de las variables cualitativas como cuantitativas, interpretándose como el efecto en la variable dependiente cuando un individuo presenta determinada modalidad respecto a aquél que presenta la categoría de referencia¹⁵.

NOTA IMPORTANTE: *En éste trabajo se tuvo la necesidad de utilizar variables explicativas cualitativas o categóricas a las cuales se les hizo el tratamiento particular descrito*

¹⁵Tomado de: <https://dialnet.unirioja.es/descarga/articulo/1281700.pdf>.

en el párrafo anterior.

Antes de indicar el proceso que se debe seguir para estimar el poder discriminante de las variables independientes, se hace necesario aclarar un conjunto de conceptos, lo cual se hace enseguida:

- **Lambda de Wilks:** Para determinar la significancia de las variables que se introducen, se emplea el estadístico Lambda de Wilks, que se obtiene de la razón entre el determinante de la matriz de varianzas y covarianzas dentro de grupos y el determinante de la matriz de varianzas y covarianzas total, que sigue una distribución F . Con esto es sencillo determinar si una variable es o no significativa para discriminar, ya que solamente basta realizar una prueba de significancia en la forma usual.
- **Distancia de Mahalanobis (D_2):** Esta técnica propuesta por Mahalanobis¹⁶ permite el cálculo de un indicador llamado distancia de Mahalanobis, que es una

¹⁶Prasanta Chandra Mahalanobis, nació el 29 de junio de 1893 y murió el 28 de junio de 1972, fue un científico de la India que destacó en estadística aplicada. Su padre, Prabodh Chandra, fue un miembro activo del movimiento religioso Bramho Samaj. Su madre, Nirodbasin, perteneció a una familia de gran tradición académica. Mahalanobis se graduó en física en 1912 por la Universidad presidencial de Kolkata, y terminó sus estudios en el College de Cambridge, para posteriormente volver a Calcuta. Fue conocido popularmente como físico por formación, estadístico por instinto y planeador por convicción, ya que usó ideas sencillas para desarrollar modelos econométricos en países en vías de desarrollo. El avizoró que la estadística, una ciencia nueva relacionada con las mediciones, tenía un amplio potencial de aplicaciones. Mahalanobis desarrolló el estadístico D_2 , conocido como o la *Distancia de Mahalanobis* (Rao, 2005). Realizó trabajos pioneros en el estudio de las variaciones antropométricas en la India, fundó el Instituto Indio de Estadística, y contribuyó al campo de las encuestas a gran escala. Inspirado por la revista científica *Biometrika* y por Acharya Brajendranath Seal, empezó a trabajar en estadística analizando resultados de exámenes universitarios, medidas antropométricas de anglo-indios de Calcuta y problemas meteorológicos. También trabajó como meteorólogo durante algún tiempo. En 1924, mientras trabajaba en la probabilidad de error de los resultados de los experimentos en agricultura, conoció a Ronald Fisher, con quien estableció una amistad que se mantendría durante toda su vida. También trabajó en modelos para prevenir inundaciones. Mahalanobis llevó a cabo tres contribuciones notables en técnicas de muestreo: proyectos piloto, diseño de proyectos óptimos e interpretación de redes de muestras. Tomado de: https://www.researchgate.net/publication/28249208_P_Ch_Mahalanobis_y_las_aplicaciones_de_su_distancia_estadistica (accessed Jul 27, 2017).

medida de la distancia entre dos puntos en el espacio, definido por dos o más variables correlacionadas. Por ejemplo, si hay dos variables no correlacionadas, pueden graficarse los puntos en un espacio de dos dimensiones, la distancia de Mahalanobis entre puntos es la distancia euclidiana. Si hay tres variables no correlacionadas puede usarse una regla para medir las distancias en el espacio de tres dimensiones para determinar la distribución entre puntos. Con más de tres variables ya no es posible graficar las distancias.

Por otro lado, cuando las variables están correlacionadas los ejes no están posicionados en ángulo recto, en esos casos la distancia euclidiana no es una medida apropiada, por tanto se usa la distancia de Mahalanobis. Para cada grupo se puede determinar la localización de un punto que representa la media para todas las variables en el espacio multivariable definido por las variables en el modelo. Estos puntos se llaman centroides de los grupos. Para cada caso puede calcularse su distancia de Mahalanobis respecto del centroide del grupo. De nuevo, puede clasificarse una observación como perteneciente al grupo al que esté más cerca o sea al cual la distancia de Mahalanobis es menor.

Teóricamente, para estimar el poder discriminante de las variables independientes se realizan los siguientes pasos:

1. Se calcula la media y la desviación estándar de cada una por grupo (0 y 1) de la muestra que se determinó para análisis. Aquellas con mayor diferencia en medias tenderán a ofrecer mayor poder. No aplica si es una variable discreta y binaria.
2. Se calcula el estadístico lambda de Wilks y el ANOVA univariante (Ratio F univariante) para valorar la significación entre las medias. Entre mayor sea el valor del Ratio F univariante se puede decir que las diferencias univariantes son más significativas.

3. Se utiliza la D^2 de Mahalanobis para llevar un proceso por etapas en la determinación de las variables con mayor capacidad discriminatoria.

En caso de no obtener resultados que permitan determinar las variables que sean significativas para discriminar, se deberá realizar un análisis factorial que permita generar nuevas variables que si pudieran serlo.

Análisis discriminante simple

En el caso de que la variable dependiente sea dicotómica, la expresión lineal de la ecuación discriminante, que representa la combinación lineal de las variables discriminantes (independientes) es la siguiente:

$$D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + e_2,$$

donde D es el punto discriminante, los β_i , con $i = 0, 1, \dots, p$ son los factores de ponderación o importancia de la variable respectiva para discriminar y X_i , con $i = 1, \dots, p$ las variables independientes o explicativas.

El proceso de modelado, consiste en encontrar el valor de los pesos que tienen las variables para discriminar, es decir, los valores de las constantes β_i , con $i = 0, 1, \dots, p$ siempre bajo el criterio de maximizar la relación entre las diferencias de los grupos con respecto a la variabilidad en los datos.

Como señalan Uriel y Aldás (2005, p-288), este criterio trata de precisar el eje discriminante de forma que las distribuciones proyectadas sobre el mismo, estén lo más separadas posibles entre sí (mayor variabilidad entre grupos) y, al mismo tiempo, que cada una de las distribuciones esté lo menos separada (menor variabilidad dentro de grupos).

Analíticamente, éste criterio puede expresarse de la siguiente manera:

$$\text{maximización de } (\lambda) = \frac{\beta' F \beta}{\beta' W \beta},$$

donde F es la matriz de covarianzas entre grupos y W calcula la matriz de la suma de cuadrados en cada grupo, sumandolos y dividiendo entre los grados de libertad. El cálculo de F será un cálculo equivalente a considerar que en cada grupo los valores de cada variable son iguales a la media de ese grupo.

La solución a éste problema se consigue derivando λ respecto de β e igualando a cero. Tras los cálculos se finaliza con un valor propio β asociado a la matriz no simétrica $W^{-1}F$. De los valores propios λ obtenidos se retiene el mayor y β es el vector característico asociado al mayor valor propio de esa matriz (Pérez, 2013).

Una vez calculada la función discriminante, si se tiene la observación de un nuevo individuo es necesario averiguar a qué grupo se le asignará y para ello hay que sustituir los valores de las variables para ese individuo en la función discriminante. El valor obtenido será la *puntuación discriminante*, que en éste caso es única al estar ante análisis discriminante simple con dos grupos. La puntuación discriminante es el resultado que se obtiene al sustituir el valor de las variables X_i , con $i = 0, 1, \dots, p$ para ese individuo en la función discriminante

$$D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p.$$

Para tomar la decisión de a qué grupo se le asigna un individuo, se crea una frontera a partir de la cual se separan los grupos. Esta frontera es llamado *punto de corte discriminante*.

Este punto de corte discriminante, es el promedio de las puntuaciones discriminantes de los denominados *centroides* de cada grupo. Los centroides son los estadísticos básicos que resumen la información sobre los grupos. Se calculan promediando

las puntuaciones discriminantes para todos los individuos dentro del grupo particular, obteniendo la media. En el caso del análisis discriminante simple, se obtienen dos centroides.

$$\bar{D}_I = (\bar{X}_{1,I}, \bar{X}_{2,I}, \dots, \bar{X}_{m,I})$$

y

$$\bar{D}_{II} = (\bar{X}_{1,II}, \bar{X}_{2,II}, \dots, \bar{X}_{m,II}).$$

Luego, para los grupos I y II se obtiene el punto de corte discriminante, así:

$$\bar{D}_I = \beta_1 \bar{X}_{1,I} + \beta_2 \bar{X}_{2,I} + \dots + \beta_m \bar{X}_{m,I}$$

$$\bar{D}_{II} = \beta_1 \bar{X}_{1,II} + \beta_2 \bar{X}_{2,II} + \dots + \beta_m \bar{X}_{m,II}$$

$$C = \frac{\bar{D}_I + \bar{D}_{II}}{2}.$$

Es así, como el criterio para clasificar a un individuo i en particular es el siguiente:

- Si $\bar{D}_I < C$, se clasifica al individuo i en el grupo I .
- Si $\bar{D}_{II} > C$, se clasifica al individuo i en el grupo II .

Al aplicarse este criterio, obviamente se van a cometer errores de clasificación (como puede verse en la Figura 3.2). El área existente a la derecha de C recoge individuos pertenecientes al grupo I , pero en los que $\bar{D}_I > C$, es decir, son individuos del grupo I mal clasificados en el grupo II . Del mismo modo, el área existente a la izquierda de C recoge individuos pertenecientes al grupo II , pero en los que $\bar{D}_{II} < C$, es decir, son individuos del grupo II mal clasificados en el grupo I (Uriel y Aldás, 2005)[4].

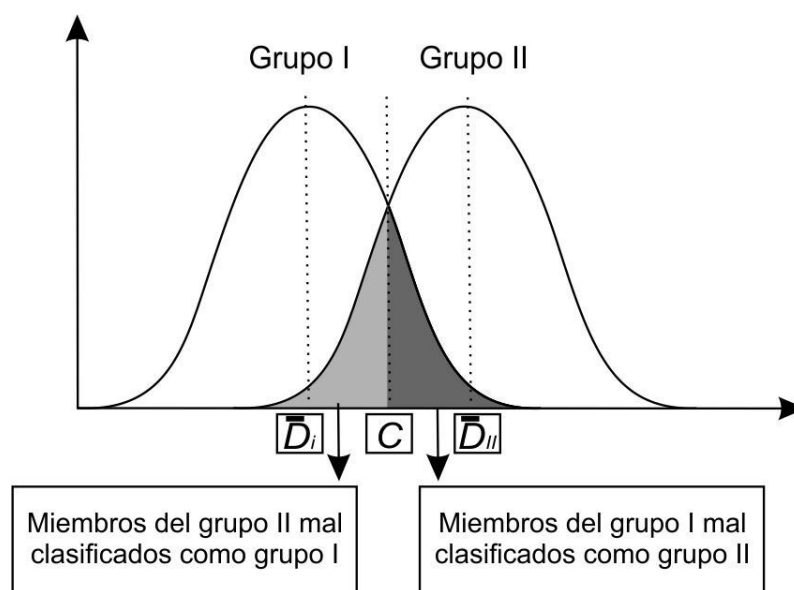


Figura 3.2: Punto de corte discriminante

3.3. Marco Normativo

3.3.1. Los Acuerdos de Basilea

La regulación financiera no es nueva en el entorno global, pues debido a la inestabilidad y la globalización de mercados financieros, la regulación es puesta en la agenda de los países desarrollados. Con la creación del *Comité de Regulación Bancaria de Basilea*, se dan los primeros pasos de la regulación internacional, cuyos lineamientos se convirtieron rápidamente en derroteros regulatorios a nivel mundial. Bajo este contexto, se hacen recomendaciones para la administración del riesgo de cartera, definido por Torres (2005) como la probabilidad de que los deudores de la entidad financiera incumplan con sus obligaciones[6].

Estos acuerdos norman las prácticas bancarias, es decir, legislan y regulan la actividad central de las entidades financieras. Estos acuerdos son emitidos por el Comité de Su-

pervisión Bancaria de Basilea y quedan como elección de cada país la incorporación de ésta normatividad al ordenamiento jurídico nacional.

Existen tres acuerdos: Basilea I, Basilea II y Basilea III.

- Basilea I (1988): El comité de Basilea hizo público su acuerdo en 1988 el cual consistió en un conjunto de recomendaciones con el objetivo de fijar un límite para el valor de los créditos que puede conceder una entidad financiera en función de su capital propio. Estas normas y recomendaciones limitaron en un 12.5% la capacidad de apalancamiento en las entidades financieras.
- Basilea II (2003): Como una falencia del acuerdo de Basilea I, quedó la falta de normatividad en cuanto a la capacidad de devolución del prestatario, es decir, el límite de crédito a conceder. En resumen, no tenía en cuenta el riesgo de crédito. Para resolver esta falencia, en 2004, el Comité de Basilea propuso un conjunto de recomendaciones, basado en tres pilares:
 - Pilar I (Los requisitos mínimos de capital): Se tiene en cuenta aquí la calidad crediticia de los prestatarios.
 - Pilar II (El proceso de la supervisión de la gestión de los fondos propios): Las entidades son responsables de mantener suficiente capital en función de sus logros, revisando los cálculos determinando el nivel de solvencia de la entidad.
 - Pilar III (La disciplina del mercado): Se establecen normas de transparencia y define la información acerca de su posición crediticia y nivel de riesgo a los mercados financieros.
- Basilea III (2010): Este acuerdo es aprobado en diciembre de 2010 e intentó adaptarse a las crisis económicas del momento. Trató de blindar a los bancos de todo el

mundo a la exposición de gran parte a los activos tóxicos en sus balances y en los derivados que circulan en el mercado. Se establecieron nuevas recomendaciones como por ejemplo:

- Endurecimiento de los criterios y aumento de la calidad del volumen de capital para asegurar su mayor capacidad para absorber pérdidas.
- Modificación de los criterios de cálculo de los riesgos para disminuir el nivel de exposición real.
- Constitución de colchones de capital durante los buenos tiempos que permitan hacer frente al cambio del ciclo económico.
- Introducción de un nuevo ratio de apalancamiento, como medida complementaria al ratio de solvencia.

Acuerdos de Basilea en Colombia

Después de la crisis internacional de los años 2007 a 2009, el *Comité de Basilea para la supervisión bancaria* emprendió, con el apoyo de la mayoría de los bancos centrales, un esfuerzo para mejorar los esquemas regulatorios y de supervisión bancaria a nivel global. En esencia, el acuerdo de Basilea III requiere más capital tangible (de mayor calidad) que Basilea I-II y, además, discrimina los requerimientos según el riesgo sistémico que encierran las entidades bancarias a nivel internacional. Existió gracias a esto un impacto en la implementación de un esquema como Basilea III en Colombia. Especialmente el fuerte impacto se debió a la sensibilidad financiera resultante de aplicar los lineamientos del Decreto 1771 de 2012. La relación de solvencia bancaria disminuiría, pasando del 14.2% observado a diciembre de 2011 a cerca de un 11.9%, bajo una hipotética aplicación inmediata de dichos lineamientos en términos de depurar el patrimonio respecto a sus componentes de impuestos diferidos netos y pasivos pen-

sionales, entre otros, y suponiendo que las entidades financieras no recomponen su capital para cumplir con esta nueva medida. Aunque a nivel general todavía se estaría cumpliendo con la solvencia mínima del 9% (e inclusive con su futuro elevamiento hacia el 10.5% que propone Basilea III), cerca de un 8.6% de los activos del sector bancario quedaría por debajo de dicho umbral, en este escenario hipotético de “ducha fría”. De allí la importancia de moverse gradualmente hacia esa depuración del capital, como bien lo contempla el Decreto 1771 de 2012, buscando cronogramas realistas que representen un equilibrio entre el norte fijado por Basilea III y la capacidad de elevar el capital tangible que tienen las entidades bancarias de Colombia.

3.3.2. Normatividad que rige a la población objeto de estudio

El segmento de ingenieros civiles contratistas SICC, son personas naturales o jurídicas, que proveen bienes y servicios a entidades públicas de forma independiente o de forma plural en consorcio y/o unión temporal mediante procesos de contratación pública y bajo la reglamentación del Estatuto General de Contratación de la Administración Pública (Ley 80 de 1993 y la Ley 1150 de 2007).

3.4. Marco Filosófico

En esta investigación se utilizaron experiencias y eventos dentro de un contexto específico dado el conocimiento que se posee por parte de los investigadores, uno como gerente con experiencia de 9 años de oficina de una entidad financiera y el otro como ingeniero civil hace que el pragmatismo sea la escuela filosófica donde se enmarca el trabajo.

3.5. Marco Situacional

En este trabajo de investigación se implementa en una entidad financiera de la ciudad de Armenia EFarmenia, un modelo scoring para el Segmento de Ingenieros Civiles Contratistas (SICC), en el Departamento del Quindío. La característica principal de esta municipalidad es que ostenta la categoría de ciudad intermedia con un desarrollo urbanístico tendiente a la alza y donde el estado es el principal empleador, motivo por el cual el subsector de los ingenieros contratistas de obras públicas cobra mucha relevancia.

La ubicación geográfica del SICC es la ciudad de Armenia, capital del departamento del Quindío. Armenia es el punto de partida para conocer un territorio que en 2011 la UNESCO declaró Patrimonio de la Humanidad: el paisaje cultural cafetero colombiano. Esta ciudad, fundada el 14 de octubre de 1889, está muy bien conectada por tierra con otros municipios donde la cultura cafetera es parte de la vida diaria de sus habitantes¹⁷. La ciudad está situada en las coordenadas 4,5 grados de latitud norte y 75,6 grados oeste a 290 kilómetros al oeste de Bogotá. Es una ciudad de tamaño medio ubicada en el triángulo de oro (Manizales, Pereira, Armenia). Su temperatura promedio está entre 18-28 grados centígrados, y es el núcleo del Área Metropolitana. Armenia acabó en el 2016 con una población de 299,712 personas. Armenia tiene una densidad de población media, de 100 habitantes por Km².

Actualmente existen 250 ingenieros civiles registrados en las bases de datos de la Cámara de Comercio de Armenia¹⁸, los cuales enfocan su principal actividad económica en la ejecución de obras públicas.

¹⁷Tomado de: <http://www.colombia.travel/es/a-donde-ir/andina/armenia>.

¹⁸Dato suministrado en noviembre de 2016 por la Cámara de Comercio de Armenia.

Por otra parte, este subsector encuentra su principal fuente de apalancamiento presupuestal en el sistema financiero, por lo cual cuentan en la ciudad con un poco más de 17 entidades bancarias para acceder a crédito para tales fines. Teniendo en cuenta lo anterior, en algunos casos, el acceso a estas líneas de crédito por parte del SICC de obras públicas, es un poco complejo, ya que esta profesión es de tipo independiente y los ingresos que se perciben son ocasionales o temporales dependiendo del tipo de labor que se desempeñe. Con base en lo anterior, en algunas ocasiones se genera un ambiente de incertidumbre para las entidades bancarias, las cuales no cuentan con una herramienta adecuada y puntual para analizar el posible riesgo de incumplimiento de este segmento del mercado. Finalmente, se hace necesario generar un modelo de scoring para otorgamiento de crédito comercial en una entidad financiera que ayude a subsanar esta situación. Para el SICC no se tienen referencias de estudios previos respecto al otorgamiento de crédito.

En este trabajo de investigación se seleccionaron únicamente los ingenieros civiles contratistas persona natural que contratan con el estado bajo la reglamentación del Estatuto General de Contratación de la Administración Pública (Ley 80 de 1993 y la Ley 1150 de 2007). Clientes de una entidad financiera de la ciudad de Armenia que de acuerdo a los servicios prestados celebran contratos bajo la modalidad de contrato de obra, para la construcción especialmente de obras de ingeniería como acueducto, alcantarillado, pavimentos, vías, puentes y todo lo relacionado con obra civil, independientemente de la modalidad de la contratación. Sin embargo, para este tipo de contrato de obra las más comunes son licitación pública y selección abreviada, y sin importar cuál sea la modalidad de ejecución y pago, estos clientes son segmentados como clientes PYME que según política de la entidad financiera sus ingresos anuales deben ser superiores a 400 millones.

DISEÑO METODOLÓGICO

El científico encuentra su recompensa en lo que Henri Poincare llama el placer de la comprensión, y no en las posibilidades de aplicación que cualquier descubrimiento pueda conllevar.

Albert Einstein

4.1. Introducción

Por los objetivos mismos planteados a nivel general y específicos, esta investigación se enmarca dentro del tipo de investigaciones de carácter descriptivo y de carácter correlacional. De carácter descriptivo porque en éste “...estudio se muestran, narran, reseñan o identifican hechos, situaciones, rasgos de un objeto de estudio, o se diseñan productos, modelos, prototipos, guías, etcétera”. De carácter correlacional porque un aspecto importante de ésta investigación consiste en “...examinar relaciones entre

variables o sus resultados, pero en ningún momento explica que una sea causa de la otra” [2].

4.2. Población o universo

La población objeto de este estudio es el segmento de ingenieros civiles contratista, con ingresos de 400 a 20000 millones de pesos anuales definidos por la entidad EFArmenia como clientes del segmento PYME.

4.3. Muestra

La muestra está conformada por los clientes de éste segmento a los que se les aprobó crédito desde el año 2012 y con corte al año 2016. La muestra consta de 101 clientes (totalidad de los clientes de este segmento de la entidad financiera EFArmenia).

4.4. Delimitación del estudio

- Espacial: Oficina de la entidad EFArmenia de la ciudad de Armenia en el departamento del Quindío.
- Demográfica: Clientes Pymes del segmento de ingenieros civiles contratistas (SICC).
- Temporal: 12 meses calendario, a partir del momento de la conceptualización y aprobación del anteproyecto.
- Temática: Creación de un modelo scoring para el otorgamiento de crédito al SICC.

4.5. Variables e indicadores del estudio

En los Cuadros 4.1 y 4.2 se presentan las variables que conforman la base de datos suministrada por el entidad EF Armenia. Allí se presentan con su respectiva descripción, clasificación, escala e indicador.

La base de datos se encuentra en el archivo digital en un CD anexo a éste trabajo.

VARIABLE	DESCRIPCIÓN	CLASIFICACIÓN	ESCALA	INDICADOR
1. Calificación Cifin (CIFÍN)	Calificación en central de riesgo Cifin en el momento del otorgamiento de crédito	Nominal dicotómica	Aprobado, rechazado	Rango de escala
2. Contratos adjudicados (CONTRATOS_ADJUDICADOS)	Total de contratos adjudicados en los últimos dos años en el momento de la aprobación del crédito	Cuantitativa intervalar	Números enteros	Número entero
3. Garantías que respaldan el crédito (GARANTIA)	Garantía ofrecida al banco para respaldar los créditos	Cualitativa nominal	Hipotecaria, FNG , Prendaria, Otras	Rango de escala
4. Nivel de endeudamiento (NE)	Nivel de endeudamiento en el momento del estudio del crédito	Cuantitativa Intervalar	Número entero	Número entero
5. Edad (EDAD)	Edad del cliente en el momento del estudio del crédito	Cuantitativa intervalar	Número entero	Número entero
6. Ingreso (INGRESO)	Ingreso anual del cliente en el último año	Cuantitativa intervalar	Número entero	Número entero
7. Género (GENERO)	Sexo biológico de pertenencia	Cualitativa nominal	Masculino, Femenino	Rango de escala

Cuadro 4.1: Variables a considerar en el modelado (1).

VARIABLE	DESCRIPCIÓN	CLASIFICACIÓN	ESCALA	INDICADOR
8. Número de entidades contratantes (EC)	Número de entidades contratantes en el histórico del cliente en el momento del estudio del crédito	Cuantitativa intervalar	Número entero	Número entero
9. Estrato socio-económico (ESTRATO)	Estrato socio-económico del cliente en el momento del estudio del crédito	Cualitativa ordinal	Número entero	Número entero
10. Número de años de experiencia en la actividad (EXPERIENCIA)	Número de años de experiencia del cliente en la actividad en el momento del estudio del crédito	Cuantitativa intervalar	Número entero	Número entero
11. Estado civil (ESTADO_CIVIL)	Estado civil del cliente en el momento del estudio del crédito	Cualitativa nominal	Soltero, Casado, Unión libre, Divorciado, viudo	Rango de escala
12. Número de hijos (HIJOS)	Número de hijos del cliente en el momento del estudio del crédito	Cuantitativa intervalar	Número entero	Número entero
13. Porcentaje otros ingresos (OTROS_INGRESOS)	Porcentaje de otros ingresos con respecto al total de ingresos	Cuantitativa intervalar	Número real	Número real
14. Tipo de vivienda (VIVIENDA)	Tipo de vivienda del cliente en el momento del estudio del crédito	Cualitativa nominal	Propia, arrendada, familiar	Rango de escala
15. Personas a cargo (PERSONAS_CARGO)	Personas a cargo	Cuantitativa intervalar	Número entero	Número entero

Cuadro 4.2: Variables a considerar en el modelado (2).

4.6. Instrumentos para recolección de la información

La información está en los registros históricos de los clientes de este segmento que durante los años 2012 a 2016 obtuvieron crédito con la entidad financiera EFArmenia. Se tabularan en un archivo Excel para respetar la petición de la empresa respecto a la confidencialidad de la información presentada por los clientes.

4.7. Procesamiento y análisis de la información

Se implementará un modelo que sea válido en aplicaciones prácticas desarrollado bajo la metodología estadística del análisis discriminante, obteniendo la posibilidad de estimar la probabilidad de incumplimiento, con el cual se puedan efectuar comparaciones de las bondades y desventajas de éste respecto al estado de comportamiento de pago de los clientes analizados de la entidad financiera EFArmenia.

Se utilizará el software Infostat para la estimación de los parámetros del modelo y para desarrollar las pruebas de bondad de ajuste, normalidad, entre otras.

Capítulo 5

MODELO DE CREDIT SCORING PARA EL SICCC

*En gran medida, la econometría no es una ciencia,
definida por un estrecho conjunto de teoremas,
sino más bien, constituye un enfoque que solo puede ser apreciado
y asimilado en toda su magnitud a través de su empleo cotidiano.*

Michael D. Intriligator

5.1. Modelación por medio de análisis discriminante

Para la ordenación de las variables y los cálculos necesarios se utilizará el software Infostat¹ en su versión con licencia profesional. La base de datos contiene el histórico de

¹InfoStat es un software para análisis estadístico de aplicación general desarrollado bajo la plataforma Windows. Tomado de: <http://www.infostat.com.ar/>.

seguimiento de 101 clientes, con fecha de corte a mayo de 2016, en donde se identifican 15 variables con relación a cada registro de clientes de la entidad. De la muestra de los 101 clientes, se tomarán 79 para realizar el modelado y los 22 restantes se dejaron para la validación del mismo.



Figura 5.1: Logo del software estadístico Infostat

5.1.1. Uso de Infostat

Una vez se tiene la base de datos en una tabla de infostat, se sigue la ruta que se observa en la Figura 5.2 para acceder a la herramienta del análisis discriminante.

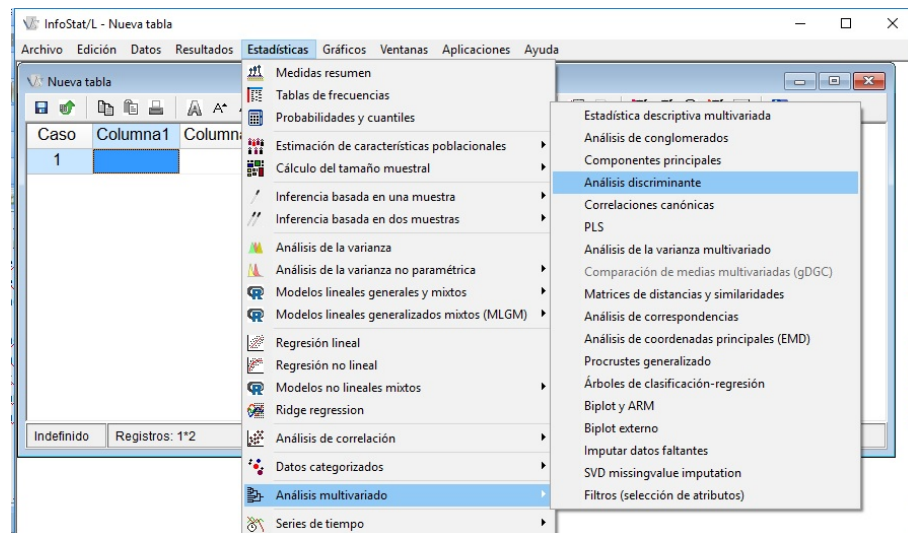


Figura 5.2: Ruta en Infostat para realizar análisis discriminante

Una vez se da click en la pestaña **Análisis Discriminante**, se obtiene la ventana tal como se observa en la Figura 5.3.

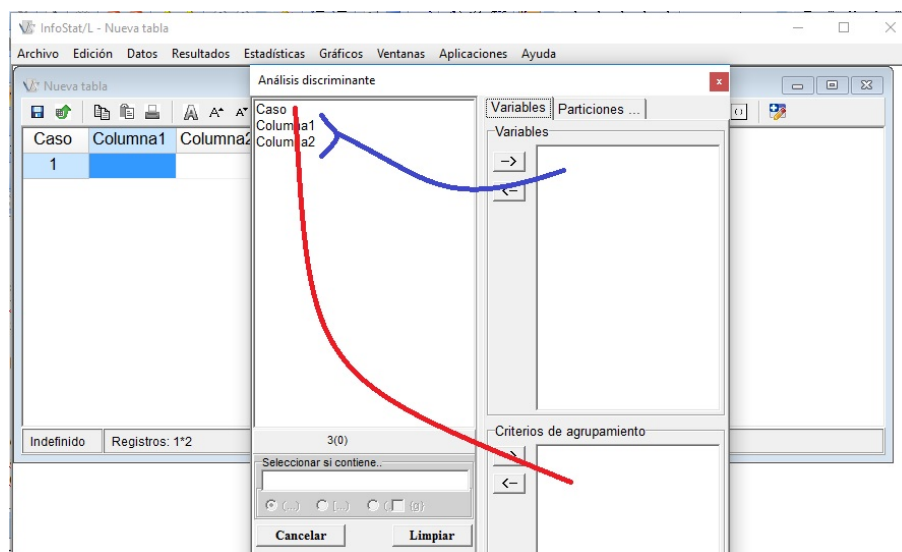


Figura 5.3: Ingreso de las variables

En éste ejemplo simple, la variable **Caso** será la que genera el criterio de agrupamiento y las variables explicativas son en éste caso **Columna1** y **Columna2**.

5.1.2. Variables

Se han seleccionado las variables con las cuales se cuenta en la base de datos suministrada por la entidad financiera EF Armenia y las cuales fueron presentadas con detalle en la Sección 4.5 del Capítulo 4. Se presenta en el Cuadro 5.1.

La variable dependiente o explicada, será **DEFAULTPAGO**, que tomará los valores 0 y 1 si el cliente es moroso (default) o pago cumplidamente (pago). La categoría de default consiste en el número de días en el cual la institución, dadas sus políticas internas, considerará que un cliente entra en mora o en incumplimiento de las obligaciones. El

# de la Variable	Notación	Descripción
1	CIFÍN	Cualitativa nominal
2	CONTRATOS_ADJUDICADOS	Cuantitativa
3	GARANTÍA	Cualitativa nominal
4	NE	Cualitativa nominal
5	EDAD	Cuantitativa
6	INGRESO	Cuantitativa
7	GÉNERO	Cualitativa nominal
8	ENTIDADES_CONTRATANTES	Cuantitativa
9	ESTRATO	Cualitativa nominal
10	EXPERIENCIA	Cuantitativa
11	ESTADO_CIVIL	Cualitativa nominal
12	HIJOS	Cuantitativa
13	OTROS_INGRESOS	Cualitativa nominal
14	VIVIENDA	Cualitativa nominal
15	PERSONAS_CARGO	Cuantitativa

Cuadro 5.1: Variables de la base de datos inicial.

default en la entidad EFArmenia de la base de datos es de 15 días, es decir, un default bueno es aquel que este por debajo de este número de días y un default malo es aquel que se encuentre por encima de éste número en mora de pago.

Las variable independientes o explicativas **determinarán** la variable dependiente del modelo propuesto.

5.2. Modelación

5.2.1. Análisis descriptivo previo

Para iniciar la aplicación del análisis discriminante se precede con un análisis descriptivo previo para poder observar inicialmente las diferencias en el comportamiento de las variables explicativas o discriminatorias en relación a la variable explicada o discriminada.

Para un análisis descriptivo inicial y para el análisis discriminante se tomará una submuestra de 79 observaciones y se dejarán las otras 22 para realizar la validación del mismo. Para tal efecto, se utilizan los datos del archivo BASEDATOS1.IDB2. También se tendrá en cuenta que para identificar relaciones de dependencia entre variables cualitativas se utiliza un contraste estadístico basado en el estadístico χ^2 (Chi-cuadrado), cuyo cálculo permitirá afirmar con un nivel de confianza estadístico determinado si los niveles de una variable cualitativa influyen en los niveles de la otra variable nominal analizada. El cálculo de la Chi-cuadrado nos permitiría saber si los valores de una variable son un factor determinante para que un cliente este en la condicion de *Default* o *Pago*.

Igualmente, se realizan estadísticas descriptivas y gráficos de dispersión de cada variable discriminadora, veamos:

Análisis de las variables cuantitativas

1. **Número de contratos adjudicados:** En las Figuras 5.4 y 5.5 se observan algunas estadísticas descriptivas de ésta variable. Se observa que en las estadísticas discriminadas según grupo al parecer hay diferencia, es decir, el grupo que está en *default* tiene una media de contratos adjudicados menor que en el caso del grupo que está en *pago*.

Variable	n	Media	D.E.	Mín	Máx	Mediana
CONTRATOS ADJUDICADOS	79	4,14	2,41	1,00	10,00	4,00

Figura 5.4: Estadísticas de la variable *Número de contratos adjudicados*.

DEFAULTPAGO	Variable	n	Media	D.E.	Mín	Máx	Mediana
0	CONTRATOS_ADJUDICADOS	33	2,30	1,24	1,00	4,00	2,00
1	CONTRATOS_ADJUDICADOS	46	5,46	2,17	2,00	10,00	5,00

Figura 5.5: Estadísticas discriminadas según grupo para la variable *Número de contratos adjudicados*.

En Análisis de varianza para ésta variable, tomando como variable dependiente DEFAULTPAGO arroja que efectivamente son diferentes, es decir, que la media del contratos de los clientes en Default es más baja que la media de contratos de los clientes en Pago. Efectivamente el p -valor es $0,0001 < 0,05$. Los resultados de éste análisis se observan en la Figura 5.6. Estos resultados dan la idea de que la variable CONTRATOS_ADJUDICADOS será de importancia en el análisis discriminante.

2. Edad:

La variable EDAD, muestra una media practicamente igual en los dos grupos. No se

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
CONTRATOS ADJUDICADOS	79	0,42	0,42	44,43

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	191,09	1	191,09	56,51	<0,0001
DEFAULTPAGO	191,09	1	191,09	56,51	<0,0001
Error	260,38	77	3,38		
Total	451,47	78			

Figura 5.6: Análisis de varianza en la variable *Número de Contratos Adjudicados*.

Variable	n	Media	D.E.	Min	Máx	Mediana
EDAD	79	46,80	7,72	35,00	60,00	46,00

Figura 5.7: Estadísticas de la variable *Edad*.

DEFAULTPAGO	Variable	n	Media	D.E.	Min	Máx	Mediana
0	EDAD	33	46,79	7,84	35,00	59,00	46,00
1	EDAD	46	46,80	7,72	35,00	60,00	45,50

Figura 5.8: Estadísticas discriminadas según grupo para la variable *Edad*.

justifica mostrar aquí el análisis de varianza, sin embargo el resultado del mismo arroja un p -valor de 0,7298, lo que confirma la igualdad de medias. Esta variable será descartada automáticamente del análisis discriminante.

3. Número de hijos:

Con la variable HIJOS, sucede algo similar que con la variable EDAD, muestra una media practicamente igual en los dos grupos. No se justifica mostrar aquí el análisis de varianza, sin embargo el resultado del mismo arroja un p -valor de 0,3677, lo que confirma la igualdad de medias. Esta variable será descartada automáticamente del análisis discriminante.

Variable	n	Media	D.E.	Mín	Máx	Mediana
HIJOS	79	1,49	1,19	0,00	3,00	2,00

Figura 5.9: Estadísticas de la variable *Hijos*.

DEFAULTPAGO	Variable	n	Media	D.E.	Mín	Máx	Mediana
0	HIJOS	33	1,45	1,09	0,00	3,00	1,00
1	HIJOS	46	1,52	1,26	0,00	3,00	2,00

Figura 5.10: Estadísticas discriminadas según grupo para la variable *Hijos*.

4. Ingresos anuales:

Resumen INGRESOS ANUALES	
n	79,00
Media	1833242236,00
D.E.	2451301758,05
Mín	234987023,00
Máx	16585997663,00
Mediana	1178511907,00

Figura 5.11: Estadísticas para la variable *Ingresos anuales*.

En Análisis de varianza para ésta variable, tomando como variable dependiente

DEFAULTPAGO	Variable	n	Media	D.E.	Mín	Máx
0	INGRESOS ANUALES	33	845111538,33	273266360,87	234987023,00	1330424665,00
1	INGRESOS ANUALES	46	2542118606,07	3022017509,72	654837129,00	16585997663,00

Figura 5.12: Estadísticas discriminadas según grupo para la variable *Ingresos anuales*.

DEFAULTPAGO arroja que efectivamente son diferentes, es decir, que la media de ingresos anuales de los clientes en Default es más baja que la media de ingresos anuales de los clientes en Pago. Efectivamente el p -valor es $0,0019 < 0,05$. Los resultados de éste análisis se observa en la figura 5.13. Estos resultados dan la idea de que la variable INGRESOS_ANUALES será de importancia en el análisis discriminante.

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
INGRESOS ANUALES	79	0,12	0,11	126,39

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	5,53365376668446E19	1	5,53365376668446E19	10,31	0,0019
DEFAULTPAGO	5,53365376668446E19	1	5,53365376668446E19	10,31	0,0019
Error	4,13356126436003E20	77	5,36826138228575E18		
Total	4,68692664102847E20	78			

Figura 5.13: Análisis de varianza para la variable *Ingresos Anuales*.

5. Número de entidades contratantes:

Variable	n	Media	D.E.	Mín	Máx	Mediana
ENTIDADES CONTRATANTES	79	2,71	1,39	1,00	5,00	2,00

Figura 5.14: Estadísticas de la variable *Número de entidades contratantes*.

En Análisis de varianza para ésta variable, tomando como variable dependiente DEFAULTPAGO arroja que efectivamente son diferentes, es decir, que la media del número de entidades contrantes que poseen los clientes en Default es

DEFAULTPAGO	Variable	n	Media	D.E.	Mín	Máx	Mediana
0	ENTIDADES_CONTRATANTES	33	1,52	0,51	1,00	2,00	2,00
1	ENTIDADES_CONTRATANTES	46	3,57	1,17	2,00	5,00	4,00

Figura 5.15: Estadísticas discriminadas según grupo para la variable *Número de entidades contratantes*.

más baja que el número de entidades contratantes que poseen los clientes en Pago. Efectivamente el p -valor es $0,0001 < 0,05$. Los resultados de éste análisis se observa en la Figura 5.16. Estos resultados dan la idea de que la variable ENTIDADES_CONTRATANTES será de importancia en el análisis discriminante.

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
ENTIDADES_CONTRATANTES	79	0,54	0,53	35,08

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo.	80,76	1	80,76	89,41	<0,0001
DEFAULTPAGO	80,76	1	80,76	89,41	<0,0001
Error	69,55	77	0,90		
Total	150,30	78			

Figura 5.16: Análisis de varianza para la variable *Número de entidades contratantes*.

6. Número de años de experiencia:

Con la variable EXPERIENCIA, sucede algo similar que con la variable EDAD, muestra una media prácticamente igual en los dos grupos. No se justifica mostrar aquí el análisis de varianza, sin embargo el resultado del mismo arroja un p -valor de 0,4592, lo que confirma la igualdad de medias. Esta variable será descartada automáticamente del análisis discriminante.

Variable	n	Media	D.E.	Mín	Máx	Mediana
EXPERIENCIA	79	15,53	5,98	5,00	25,00	16,00

Figura 5.17: Estadísticas para la variable *Número de años de experiencia*.

7. Número de personas a cargo:

DEFAULTPAGO	Variable	n	Media	D.E.	Min	Máx	Mediana
0	EXPERIENCIA	33	14,94	6,23	5,00	25,00	18,00
1	EXPERIENCIA	46	15,96	5,82	5,00	25,00	16,00

Figura 5.18: Estadísticas según grupo para la variable *Número de años de experiencia*.

Con la variable PERSONAS_CARGO, sucede algo similar que con la variable EDAD, muestra una media prácticamente igual en los dos grupos. No se justifica mostrar aquí el análisis de varianza, sin embargo el resultado del mismo arroja un p -valor de 0,4592, lo que confirma la igualdad de medias. Esta variable será descartada automáticamente del análisis discriminante.

Variable	n	Media	D.E.	Min	Máx	Mediana
PERSONAS_CARGO	79	2,22	1,23	0,00	4,00	2,00

Figura 5.19: Estadísticas para la variable *Número de personas a cargo*.

DEFAULTPAGO	Variable	n	Media	D.E.	Min	Máx	Mediana
0	PERSONAS_CARGO	33	2,30	1,10	0,00	4,00	2,00
1	PERSONAS_CARGO	46	2,15	1,32	0,00	4,00	2,00

Figura 5.20: Estadísticas según grupo para la variable *Número de personas a cargo*.

Análisis de las variables cualitativas

Para éste conjunto, se analizan dos variables (que admiten distintas modalidades) mediante una tabla de contingencia. Cuando una muestra aleatoria que se obtiene de una población se clasifica en dos rangos diferentes, en los que cada rango se encuentra subdividido en cierto número de categorías, el resultado recibe el nombre de *tabla de contingencia con dos criterios de clasificación* o de doble entrada. Esta tabla se forma por las frecuencias relativas que se observaron para las dos categorías. La intersección

entre una fila y una columna da lugar a una celda o casilla, cuya frecuencia observada es n_{ij} . A pesar de que en éste trabajo sólo se analizarán tablas de contingencia para dos clasificaciones, es posible analizar tablas que contengan más de dos clasificaciones. El análisis de una tabla de contingencia de doble entrada supone que las dos clasificaciones son independientes. Es decir, bajo la hipótesis nula de independencia se desea determinar si existe una suficiente diferencia entre las frecuencias que se observan y las correspondientes frecuencias que se esperan, tal que la hipótesis nula se rechace. Se contrasta la hipótesis nula que presupone la independencia entre ambas variables, mediante el estadístico χ^2 de Pearson que permite contrastar las hipótesis:

$$\begin{cases} H_0: \text{Ambas variables son independientes} \\ H_1: \text{Existe una relación de dependencia} \end{cases}$$

En éste caso una de las variables será la que se tiene por variable respuesta, es decir, DEFAULTPAGO.

1. **Calificación Cifín:** Como se puede observar en la Figura 5.21, el 95% de los clientes poseen calificación Cifín aprobada. El 5% restante, que corresponde a solo 4 clientes son casos excepcionales aprobados internamente por el gerente de oficina o el gerente de zona. Es una variable que no discrimina en absoluto y será descartada del análisis discriminante final.
2. **Género:** Con ésta variable pasa algo muy similar que con la variable CIF IN. Como se puede observar en la Figura 5.21, el 97% de los clientes son de género masculino. El 3% restante, que corresponde a solo 2 clientes son pertenecientes al género femenino. Es una variable que no discrimina en absoluto y será descartada del análisis discriminante final.

Variable	Clase	MC	FA	FR
CIFÍN	1	0	75	0,95
CIFÍN	2	1	4	0,05

Variable	Clase	MC	FA	FR
GENERO	1	0	2	0,03
GENERO	2	1	77	0,97

Variable	Clase	MC	FA	FR
GARANTÍA	1	0	21	0,27
GARANTÍA	2	1	58	0,73

Variable	Clase	MC	FA	FR
VIVIENDA	1	0	10	0,13
VIVIENDA	2	1	69	0,87

Variable	Clase	MC	FA	FR
ESTADO_CIVIL	1	0	65	0,82
ESTADO CIVIL	2	1	14	0,18

Variable	Clase	MC	FA	FR
OTROS_INGRESOS	1	0	69	0,87
OTROS INGRESOS	2	1	10	0,13

Variable	Clase	Categorías	FA	FR
NE	1	ALTO	12	0,15
NE	2	BAJO	58	0,73
NE	3	MEDIO	9	0,11

Figura 5.21: Estadísticas para la variables: *Garantía, Nivel de endeudamiento, Género, Estrato, Estado civil, Tipo de vivienda, Otros ingresos y Calificación Cifín.*

3. **Garantía:** Como se puede observar en la Figura 5.21, el 73% de los clientes han presentado garantía al banco. El 27% restante, obtuvieron su crédito sin garantía. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.26.
4. **Tipo de vivienda:** Se puede observar en la Figura 5.21, que el 87% de los clientes poseen vivienda propia. El 13% restante, poseen vivienda en arriendo. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.27.
5. **Estado civil:** Se puede observar en la Figura 5.21, que el 82% de los clientes se encuentran casados. El 18% restante, se encuentran solteras. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.28.
6. **Otros ingresos:** Se puede observar en la Figura 5.21, que el 87% de los clientes certifican otros ingresos. El 13% restante, no registran otros ingresos. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.29.
7. **Nivel de endeudamiento:** En la Figura 5.21, el 15% de los clientes poseen un alto nivel de endeudamiento. El 73% poseen un nivel medio de endeudamiento y el 11% restante, poseen un bajo nivel de endeudamiento. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.30.
8. **Estrato socioeconómico:** En la Figura 5.22, se tiene la variable ESTRATO renombrada como CATESTRATO. En ella se puede observar que el 43% de los clientes están ubicados en estrato 4, el 29% ubicados en estrato 5 y el 28% restante, están ubicados en estrato 6. Se realiza para éste caso un análisis por medio de tablas de contingencia. Obsérvese la Figura 5.23.

Al observar la primera tabla en las Figuras 5.24 a ??, si no hay diferencias de respuestas,

Variable	Clase	Categorías	FA	FR
CATESTRATO	1	Estrato4	34	0,43
CATESTRATO	2	Estrato5	23	0,29
CATESTRATO	3	Estrato6	22	0,28

Figura 5.22: Variable ESTRATO renombrada como CATESTRATO.

En columnas:ESTRATO

DEFAULTPAGO	4,00	5,00	6,00	Total
0	16	10	7	33
1	18	13	15	46
Total	34	23	22	79

Frecuencias esperadas bajo independencia

En columnas:ESTRATO

DEFAULTPAGO	4,00	5,00	6,00	Total
0	14,20	9,61	9,19	33,00
1	19,80	13,39	12,81	46,00
Total	34,00	23,00	22,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	1,31	2	0,5183
Chi Cuadrado MV-G2	1,34	2	0,5122
Coef.Conting.Cramer	0,09		
Coef.Conting.Pearson	0,13		

Figura 5.23: Estadísticas variable ESTRATO.

En columnas:CIFÍN

DEFAULTPAGO	0	1	Total
0	31	2	33
1	44	2	46
Total	75	4	79

Frecuencias esperadas bajo independencia
En columnas:CIFÍN

DEFAULTPAGO	0	1	Total
0	31,33	1,67	33,00
1	43,67	2,33	46,00
Total	75,00	4,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	0,12	1	0,7320
Chi Cuadrado MV-G2	0,12	1	0,7337
Irwin-Fisher bilateral	-0,02		>0,9999
Coef.Conting.Cramer	0,03		
Kappa (Cohen)	-0,01		
Coef.Conting.Pearson	0,04		
Coefficiente Phi	-0,04		

Figura 5.24: Tabla de contingencia variables CIF IN vs PAGODEFAULT

En columnas:GENERO

DEFAULTPAGO	0	1	Total
0	1	32	33
1	1	45	46
Total	2	77	79

Frecuencias esperadas bajo independencia
En columnas:GENERO

DEFAULTPAGO	0	1	Total
0	0,84	32,16	33,00
1	1,16	44,84	46,00
Total	2,00	77,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	0,06	1	0,8111
Chi Cuadrado MV-G2	0,06	1	0,8124
Irwin-Fisher bilateral	0,01		>0,9999
Coef.Conting.Cramer	0,02		
Kappa (Cohen)	0,01		
Coef.Conting.Pearson	0,03		
Coefficiente Phi	0,03		

Figura 5.25: Tabla de contingencia variables GENERO vs PAGODEFAULT

En columnas: GARANTÍA

DEFAULTPAGO	0	1	Total
0	7	26	33
1	14	32	46
Total	21	58	79

Frecuencias esperadas bajo independencia

En columnas: GARANTÍA

DEFAULTPAGO	0	1	Total
0	8,77	24,23	33,00
1	12,23	33,77	46,00
Total	21,00	58,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	0,84	1	0,3601
Chi Cuadrado MV-G2	0,85	1	0,3562
Irwin-Fisher bilateral	-0,09		0,4431
Coef. Conting. Cramer	0,07		
Kappa (Cohen)	-0,10		
Coef. Conting. Pearson	0,10		
Coefficiente Phi	-0,10		

Figura 5.26: Tabla de contingencia variables GARANTIA vs PAGODEFAULT

En columnas: VIVIENDA

DEFAULTPAGO	0	1	Total
0	0	33	33
1	10	36	46
Total	10	69	79

Frecuencias esperadas bajo independencia

En columnas: VIVIENDA

DEFAULTPAGO	0	1	Total
0	4,18	28,82	33,00
1	5,82	40,18	46,00
Total	10,00	69,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	8,21	1	0,0042
Chi Cuadrado MV-G2	11,84	1	0,0006
Irwin-Fisher bilateral	-0,22		0,0041
Coef. Conting. Cramer	0,23		
Kappa (Cohen)	-0,24		
Coef. Conting. Pearson	0,31		
Coefficiente Phi	-0,32		

Figura 5.27: Tabla de contingencia variables VIVIENDA vs PAGODEFAULT

En columnas:ESTADO_CIVIL

DEFAULTPAGO	0	1	Total
0	29	4	33
1	36	10	46
Total	65	14	79

Frecuencias esperadas bajo independencia

En columnas:ESTADO_CIVIL

DEFAULTPAGO	0	1	Total
0	27,15	5,85	33,00
1	37,85	8,15	46,00
Total	65,00	14,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	1,22	1	0,2695
Chi Cuadrado MV-G2	1,26	1	0,2611
Irwin-Fisher bilateral	0,10		0,3738
Coef.Conting.Cramer	0,09		
Kappa (Cohen)	0,08		
Coef.Conting.Pearson	0,12		
Coefficiente Phi	0,12		

Figura 5.28: Tabla de contingencia variables ESTADO_CIVIL vs PAGODEFAULT

En columnas:OTROS_INGRESOS

DEFAULTPAGO	0	1	Total
0	33	0	33
1	36	10	46
Total	69	10	79

Frecuencias esperadas bajo independencia

En columnas:OTROS_INGRESOS

DEFAULTPAGO	0	1	Total
0	28,82	4,18	33,00
1	40,18	5,82	46,00
Total	69,00	10,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	8,21	1	0,0042
Chi Cuadrado MV-G2	11,84	1	0,0006
Irwin-Fisher bilateral	0,22		0,0041
Coef.Conting.Cramer	0,23		
Kappa (Cohen)	0,19		
Coef.Conting.Pearson	0,31		
Coefficiente Phi	0,32		

Figura 5.29: Tabla de contingencia variables OTROS_INGRESOS vs PAGODEFAULT

En columnas:NE				
DEFAULTPAGO	ALTO	BAJO	MEDIO	Total
0	3	26	4	33
1	9	32	5	46
Total	12	58	9	79

Frecuencias esperadas bajo independencia				
En columnas:NE				
DEFAULTPAGO	ALTO	BAJO	MEDIO	Total
0	5,01	24,23	3,76	33,00
1	6,99	33,77	5,24	46,00
Total	12,00	58,00	9,00	79,00

Estadístico	Valor	gl	p
Chi Cuadrado Pearson	1,64	2	0,4411
Chi Cuadrado MV-G2	1,72	2	0,4224
Coef.Conting.Cramer	0,10		
Coef.Conting.Pearson	0,14		

Figura 5.30: Tabla de contingencia variables NE vs PAGODEFAULT

es decir, si las frecuencias absolutas en una misma fila son muy similares es que no hay relación entre las variables y se sospecha que las variables son independientes; y en caso contrario, es decir, si hay diferencias de respuestas, es decir, si las frecuencias absolutas en una misma fila no son similares es que hay relación entre las variables y se sospecha que las variables son dependientes. Esta primera tabla (en cada figura), se compara con la segunda tabla (en cada figura) que muestra *las frecuencias esperadas bajo independencia*. En otras palabras, si las primeras dos tablas (de cada figura) son similares, se sospecha independencia en las variables respectivas.

Analizando la tercera tabla (en cada una de la las Figuras 5.24 a ??) donde se observa entre otras cosas el *estadístico Chi-cuadrado de Pearson* y el *estadístico chi-cuadrado MV-G2*, se puede determinar cuáles variables son dependientes y cuáles son independientes con respecto a la variable dependiente PAGODEFAULT.

En resumen, se presenta en el Cuadro 5.2 los respectivos estadísticos para cada caso representado en las Figuras 5.24 a 5.30. Se puede observar que se han marcado con *

aquellas variables para las cuales se ha aceptado la hipótesis nula de independencia. Estas variables no se tendrán en cuenta para el análisis discriminante dado que no se ha podido rechazar la hipótesis.

ESTADÍSTICOS DE PRUEBA		
	ESTADÍSTICOS DE PRUEBA	
	Chi-cuadrado de Pearson	Chi-cuadrado MV-G2
CIFIN*	0,7320	0,7337
GENERO*	0,811	0,8124
GARANTIA*	0,3601	0,3562
VIVIENDA	0,0042	0,0006
ESTADO_CIVIL*	0,2695	0,2611
OTROS_INGRESOS	0,0042	0,0006
NE*	0,4411	0,4224
ESTRATO*	0,5183	0,5122

Cuadro 5.2: Estadísticos Chi-cuadrado de Pearson y MV-G2

5.2.2. Variables para el análisis discriminante

Una vez realizado el análisis descriptivo previo de la sección 5.2.1, finalmente se han eliminado algunas variables por no presentar dependencia o alguna relación estadística con la variable PAGODEFAULT. Ellas se presentan en el Cuadro 5.3:

5.3. Proceso de modelación

Con las cinco variables que se tiene, se realiza el ejercicio de modelar por medio análisis discriminante y con la ayuda del software Infostat. En el caso que corresponde a éste trabajo, se desea poder predecir el riesgo de adjudicar crédito a un individuo del

VARIABLES			
1	Número de contratos adjudicados	CONTRATOS_ADJUDICADOS	x_1
2	Ingresos anuales	INGRESOS_ANUALES	x_2
3	Número de entidades contratantes	ENTIDADES_CONTRATANTES	x_3
4	Otros ingresos	OTROS_INGRESOS	x_4
5	Tipo de Vivienda	VIVIENDA	x_5

Cuadro 5.3: Variables a tener en cuenta en la aplicación del Análisis discriminante

SICC. Para tal efecto, se acude el archivo o fichero BASEDEDATOS.idb2 que consta de 79 clientes de la EF Armenia, escogidos aleatoriamente de una base total de 101. La variables potencialmente explicativas son: número de contratos adjudicados, ingresos anuales, número de entidades contratantes, tipo de vivienda del cliente y otros ingresos.

Para iniciar el análisis se procede como se indicó en la Figura 5.2 donde se introduce como criterio de agrupación a la variable dependiente PAGODEFAULT la cual ya se encuentra categorizada como 0 =no pago o default y 1 =pago.

Como variables independientes se ingresan las cinco que ya fueron determinadas. Véase la Figura 5.31. Como se puede observar se ha obtenido por el método de *introducir todas las dependientes juntas*.

Al aceptar, aparece otra ventana de diálogo donde se solicitaron las opciones (ver Figura 5.32): *Guardar clasificación*, *Mostrar tasa de error de clasificaciones*, *Guardar primeras coordenadas discriminantes* y *Sobreescribir coordenadas discriminantes*. Finalmente se selecciona *Gráfico*. Al aceptar se obtuvieron los siguientes resultados.

Al aceptar se obtuvieron los siguientes resultados (ver Figuras 5.33 y 5.34):

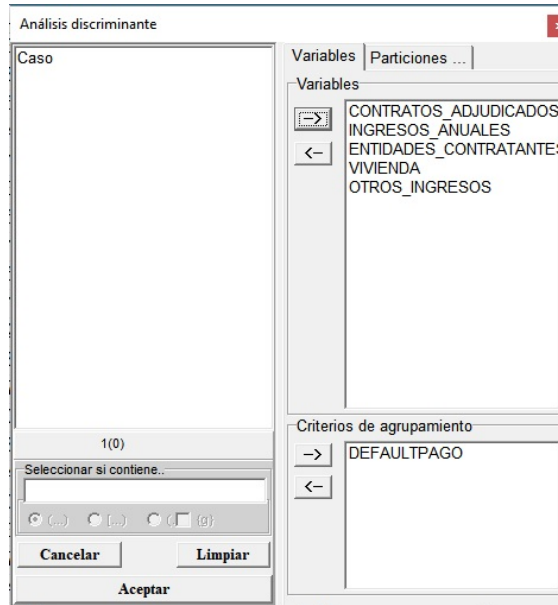


Figura 5.31: Ingreso de variables a las aplicación Infostat para modelado

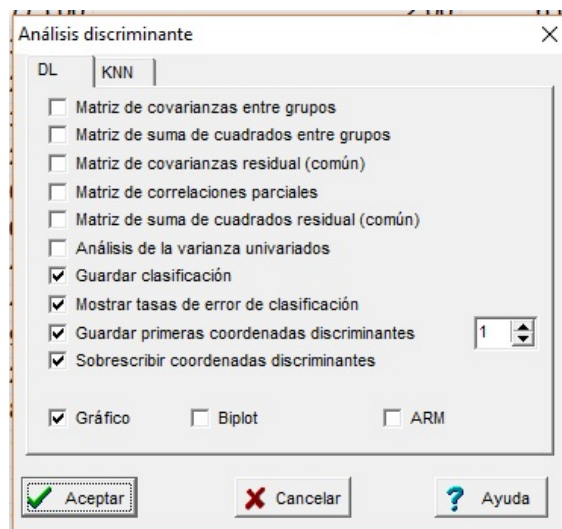


Figura 5.32: Opciones escogidas para el análisis

```

Análisis discriminante

Análisis discriminante lineal

Casos leídos 82

Variables

CONTRATOS_ADJUDICADOS
INGRESOS_ANUALES
ENTIDADES_CONTRATANTES
VIVIENDA
OTROS_INGRESOS

Variables de clasificación

DEFAULTPAGO

```

Figura 5.33: Salida del programa Infostat para análisis discriminante (I).

```

Prueba de Homogeneidad de Matrices de Covarianzas
Grupos N Estadístico gl p-valor
-----
2 79 137,48 36 <0,0001

Autovalores de Inv(E)H
Autovalores % % acumulado
-----
1,46 100,00 100,00

Funciones discriminantes canónicas
1
-----
Constante -2,52
CONTRATOS_ADJUDICADOS 0,07
INGRESOS_ANUALES 6,2E-11
ENTIDADES_CONTRATANTES 0,93
VIVIENDA -0,55
OTROS_INGRESOS 0,55

Funciones discriminantes - datos estandarizados con las varianzas comunes
1
-----
CONTRATOS_ADJUDICADOS 0,12
INGRESOS_ANUALES 0,15
ENTIDADES_CONTRATANTES 0,88
VIVIENDA -0,18
OTROS_INGRESOS 0,18

Centroides en el espacio discriminante
Grupo Eje 1
-----
0 -1,41
1 1,01

Tabla de clasificación cruzada (tasa de error aparente)
Grupo 0 1 Total Error(%)
-----
0 33 0 33 0,00
1 2 44 46 0,04
Total 35 44 79 0,03

```

Figura 5.34: Salida del programa Infostat para análisis discriminante (II).

5.3.1. Interpretación de resultados

Se analiza a continuación la salida generada por el programa Infostat según las variables incluidas y las opciones ejecutadas.

1. La prueba de homogeneidad de matrices de covarianza arrojó un valor de $p < 0,001$, sugiriendo que no se cumple este supuesto y que una función discriminante cuadrática podría ser mejor. Sin embargo, se continuó con el análisis ya que por experiencia éste puede obviarse sin prejuicio para la efectividad del modelo.
2. A partir de los autovalores de la expresión $\mathbf{inv(E)H}$, se puede concluir que el eje canónico explica el 100% de la variación entre grupos. Recuerdese que como hay sólo dos grupos, existe sólo una función discriminante, es decir un solo eje canónico.
3. La función discriminante canónica puede expresarse de la siguiente manera según el orden de las variables dado en el Cuadro 5.3:

$$D = \text{Constante} + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5,$$

de tal manera que usando los coeficientes canónicos se tiene el modelo discriminante así:

$$D = -2,52 + 0,07x_1 + 6,2 \times 10^{-11}x_2 + 0,93x_3 + 0,55x_4 - 0,55x_5.$$

4. En ésta función lineal de las cinco variables seleccionadas, los coeficientes responden a las distribuciones de cada variable. Si las variables tienen varianzas muy distintas y/o existe alta covariación entre pares de variables, la interpretación puede ser engañosa, por tal razón es conveniente analizar la importancia relativa de cada variable, en la discriminación de los grupos, usando la función de coeficientes estandarizados por varianzas y covarianzas. Con la *Función*

discriminante estandarizada por las covarianzas comunes puede observarse que ENTIDADES_CONTRATANTES (Número de entidades contratantes) es la variable mas importante para la discriminación: observaciones con valores altos aparecerán a la derecha del gráfico de dispersión de observaciones en el espacio discriminante (espacio formado por el eje canónico) ya que el coeficiente es positivo (0,88).

5. Respecto a los *centroides en el espacio discriminante* o medias de las funciones por grupo, se tiene los valores $-1,41$ para el grupo DEFAULT (0) y $1,01$ para el grupo PAGO (1). De ésta manera el valor discriminante es:

$$D_1 = \frac{(\text{Centroide Grupo 0}) - (\text{Centroide Grupo 1})}{2} = \frac{-1,41 + 1,01}{2} = -0,20$$

Los centros de gravedad o centroides (vector de medias) resumen la información sobre los grupos.

6. La *Tabla de clasificación cruzada*, señala que las 33 observaciones del grupo 0 fueron correctamente asignadas, la tasa de error en éste grupo fue del 0%. De los 46 individuos del grupo 1, 44 fueron bien asignados y solamente 2 fueron mal asignados, generando una tasa de error de solo 0,04%. El error global del modelo ahora es solo de 0,03%.

5.3.2. Validación: Back Testing

Una vez se ha realizado el ejercicio completo de determinar las variables que verdaderamente son discriminantes, se llegó al modelo final con 5 variables explicativas o discriminantes como se observa en el Cuadro 5.3 y la función discriminante canónica puede expresarse de la siguiente manera:

$$D = -2,52 + 0,07x_1 + 6,2 \times 10^{-11}x_2 + 0,93x_3 + 0,55x_4 - 0,55x_5. \quad (5.1)$$

El objetivo del Back Testing² es comprobar el ajuste y la consistencia del modelo, dado que se modeló con una parte de los datos, con la proporción sobrante o no modelada (22 datos), se realizan estimaciones y se comprueba que efectivamente los resultados predichos en el modelo se ajustan a los resultados reales.

Usando el modelo 5.1, se estimarán los scoring de los 22 datos (clientes) no modelados, para demostrar que en la realidad el modelo es consistente. Dentro de los datos no modelados, escogidos aleatoriamente, se tienen 9 registros que están en default, y los 11 registros restantes no lo están.

Luego de calculados los scoring de los 22 clientes se obtuvieron los datos que aparecen en el Cuadro

Back Testing EFArmenia				
	Resultado real			
GRUPO	Default	Pago	Total	Error %
Default	9	0	9	0,00
Pago	1	12	13	0,08
Total	10	12	22	0,04

Cuadro 5.4: Back testing de la validación del modelo

Lo que se puede concluir a partir del Cuadro 5.4, es que como era de esperar, el modelo logró predecir los créditos con alta probabilidad de default (error del 0,00%), créditos a clientes que efectivamente entraron en default, 9 en total. Hay 1 registros donde el modelo falla para asignar el puntaje a un cliente que si paga (error del 0,09%). Y final-

²Se refiere a una metodología que se utiliza para verificar y diagnosticar la eficiencia y los resultados de un modelo en particular. Esta metodología intenta determinar si las conclusiones del modelo o los cálculos de las variables estimadas son acertadas. Como su nombre lo indica, backtesting significa evaluar hacia atrás, por esta razón, el principal objetivo es saber cuáles habrían sido las conclusiones de un modelo en periodos anteriores y compararlo con datos históricos. Es posible definir de forma sencilla el backtesting, como un test de un sistema (sea cual sea) con los datos del pasado, en el que se simula de forma exacta lo que dicen las reglas de entrada y salida. Tomado de: www.rankia.co/blog/analisis-colcap/3606172-backtesting-definicion-ejemplos

mente, de los 22 registros que el modelo analiza, solo se llega a un error de 0,045%.

5.3.3. Análisis cualitativo de resultados

Con un modelo scoring como 5.1, se podrá hacer un análisis de gran valor en beneficio de la entidad financiera, para observar el comportamiento actual y a mediano plazo de sus clientes. Por ejemplo, uno de los resultados importantes es que variables como edad, número de hijos, estado civil y género no son determinantes para la elaboración del modelo dado que no discriminan las dos poblaciones en estudio. Otro resultado interesante, es que los ingresos anuales son importantes. Su coeficiente aparenta ser pequeño ($6,2 \times 10^{-11}$), pero dado la magnitud de los datos de la variable termina siendo significativo. Hacen parte del modelo variables poco intuitivas como el tipo vivienda del cliente. Esta variable tiene un coeficiente no despreciable de $(-0,55)$.

Los Ingenieros Civiles Contratistas, son sin lugar a dudas un segmento con características muy particulares dentro de las PYMES. Las variables que los identifican como por ejemplo el *número de entidades contratantes* no es comunmente analizada por los modelos clásicos de otorgamiento de crédito.

CONCLUSIONES

- Desde el principio del trabajo, quedó claro que el sistema financiero dada su inestabilidad debe ser regulado y la Superintendencia Financiera de Colombia reconoce que estas instituciones se enfrentan, entre los principales riesgos al de mercado, al operativo y al de crédito. Por eso, implementó un SARC, todo un aparato de evaluación estadística, de políticas corporativas y acumulación de información, para gestionar conjuntamente, al lado de la institución financiera, el riesgo. El activo de una entidad financiera, está representado en los créditos otorgados y estos están sujetos al riesgo de la probabilidad de impago, si la probabilidad de que este dinero no sea devuelto es alta, las instituciones financieras no tendrán la solidez que se requiere que tengan para impulsar la economía. El modelo propuesto disminuye en gran manera, desde el otorgamiento, la exposición de las entidades financieras al riesgo de cartera, asegurando desde el comienzo la solidez de este importante sector. Como lo expone Schreiner (2002), un modelo de Scoring es perfecto para las entidades financieras de países en desarrollo, pues estas no cuentan con grandes volúmenes históricos de créditos, como si los tienen los

grandes bancos de países desarrollados. Además, es un modelo de construcción simple, pero de un gran valor analítico.

- Los modelos estimados por medio del análisis discriminante son bastante precisos, como se pudo comprobar con los estadísticos calculados para valorar su significación. Adicionalmente, poseen una gran precisión en el ejercicio de la predicción. El modelo generado predice o detecta especialmente muy bien a los clientes morosos con una tasa de acierto del 99,95 % comparada con la tasa de acierto de EFArmenia que estuvo en solo el 58,23 %.
- Debe quedar claro, que un modelo como el presentado no debe reemplazar completamente a los análisis y políticas actuales, pues éstos son indispensable para la recolección de información y observación de variables que pueden no estar dentro del modelo, así como para la validación del mismo y los estudios financieros. También es importante tener en cuenta que el analista en ningún caso debe conocer las variables que en mayor medida aportan para que se otorgue el crédito, pues éste podrá acomodar variables a su tamaño para obtener un resultado. Quien calibrará, ajustará y evaluará el modelo, deberá ser un consultor que nada tendrá que ver con la recolección de información del prestatario, esto con el fin de que los analistas sólo conozcan el resultado del modelo, más no la ecuación que determina el Scoring.
- Otro aporte importante de este tipo de modelos, es que hace un análisis individual, a diferencia de otras herramientas que miden el riesgo como son los modelos de cartera y los VaR marginales. Además, el modelo de Scoring emplea un alto número de variables, en donde relaciona el riesgo dependiendo de las características del cliente y reduce el tiempo en la asignación de un crédito. Dentro de los puntos en contra, podrían destacarse que la información de cada cliente debe ser abundante y de calidad, por lo que la información debe seguir siendo recogida,

para hacer cada vez más exacta la predicción. Este tipo de modelo además supone que el futuro será igual al pasado, por lo que una estimación en un entorno económico saludable, puede no ser confiable en un entorno económico en crisis.

- Las variables que determinan que un cliente sea moroso, han resultado significativas tanto las de carácter técnico, como aquellas que reflejan el perfil del individuo.
- La conclusión global que se obtiene es que la técnica de análisis discriminante es adecuada para el estudio y predicción de clientes morosos y no morosos, consiguiendo una elevada eficacia predictiva. Esto se cumple, dado que las condiciones exigidas para la aplicación del análisis discriminante, en cuanto a la igualdad de varianzas en los dos grupos definidos por la variable dependiente y la normalidad de las variables independientes en la población de la que procede la muestra, se verifican en la muestra de clientes objeto de estudio. Por ello, los resultados obtenidos mediante esa técnica son similares a los conseguidos mediante regresión logística, que es el método alternativo aplicado para el análisis de variables cualitativas cuando se incumplen tales restricciones.
- El trabajo revisó y buscó clarificar distintos aspectos asociados a los modelos de credit scoring. Aunque nada impide que estas técnicas se apliquen a grandes empresas, y de hecho existen algunos desarrollos para ellas, los modelos de credit scoring se emplean mayormente para evaluar individuos y pequeñas y medianas empresas, mientras que las grandes empresas se analizan a través de sistemas de rating. Si bien su aplicación más conocida es en la originación de financiaciones, las entidades financieras emplean estas herramientas también con otros propósitos, como por ejemplo en el diseño de estrategias de marketing para ofrecer productos de manera proactiva y masiva.
- Los modelos de credit scoring difieren en la información que emplean, en función

del sujeto evaluado (individuos o PyMEs) y del desarrollador del modelo (con datos de buró o de la entidad financiera). En cualquier caso, la revisión de la literatura especializada muestra que en su construcción predominan los enfoques econométricos (regresión logística, modelos probit) y desarrollos de inteligencia artificial, como los árboles de decisión. Los motivos para su predominio son básicamente dos: en general las metodologías relevadas muestran resultados similares, por lo que tienden a emplearse aquellas cuyo funcionamiento e interpretación son más sencillos, en contraposición a enfoques más sofisticados y de difícil interpretación, como ser las redes neuronales. Los resultados del modelo estimado con datos de la CENDEU, construido con el sólo propósito de funcionar como ejemplo, muestran la relevancia de las siguientes variables para predecir el comportamiento de los deudores: la historia de pagos y el comportamiento corriente en otras entidades financieras; la “búsqueda de financiamiento”, aproximada por la cantidad de entidades con la que los deudores operan; si la deuda está cubierta con garantías preferidas (como los hipotecarios y prendarios) y el grado de cobertura. Otra variable que tiene una importancia significativa, y cuya inclusión en cualquier modelo es crítica, es el nivel de actividad o PBI. Su rol en modelo estimado es el esperado: en los años buenos la creación de empleo y el aumento en las ventas mejoran la calidad crediticia de individuos y PyMEs, por lo que su riesgo disminuye.

- Queda abierta la posibilidad de evaluar otros segmentos que posiblemente ameriten un modelo particular para la medición de la probabilidad de incumplimiento.

Capítulo **7**

BIBLIOGRAFÍA

Bibliografía

- [1] Asobancaria. <http://www.asobancaria.com>
- [2] Bernal, César Augusto. Metodología de la Investigación. Pearson Educación, México, 2006.
- [3] Gutiérrez Girault, Matías Alfredo. Modelos de Credit Scoring - Qué, Cómo, Cuándo y Para Qué -. Banco Central de la República Argentina (BCRA). Octubre de 2007.
- [4] Iñigo De Diego Lago, Santiago. Análisis Discriminante: Aplicación a la búsqueda de variables relacionadas con la capacidad de exportar. Trabajo de Grado en Economía, Universidad de la Coruña. España, 2005.
- [5] Muñoz Salas, Evelyn. Técnica de Análisis Discriminante: Una Aplicación para el área bancaria. Banco Central de Costa Rica, División Económica, Departamento de Investigaciones Económicas. DIE-NT-03-98, Agosto , 1998.
- [6] Ochoa, Juan Camilo. Galeano, Wilinton. Agudelo, Luis Gabriel. Construcción de un modelo scoring para el otorgamiento de crédito de una entidad financiera. Perfil de coyuntura económica No. 16, diciembre 2010, P 191-222. Universidad de Antioquía.

- [7] Portal único de Contratación. <https://www.contratos.gov.co>

- [8] Rayo Cantón, Salvador. Lara Rubio, Juan. Camino Blasco, David. Un Modelo de Credit Scoring para instituciones de microfinanzas en el marco de Basilea II. Tomado de: <http://www.esan.edu.pe/publicaciones/2010/06/02/05.pdf>

- [9] Schreiner, Mark. Ventajas y Desventajas del Scoring Estadístico para las Microfinanzas. 11 de Septiembre de 2002. Microfinance Risk Management 6970 Chipewa St. #1W, St. Louis, MO 63109-3060, U.S.A. <http://www.microfinance.com>. Washington University in St. Louis.

- [10] Superintendencia Financiera de Colombia. <http://www.superfinanciera.gov.co>