



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

**FACULTAD DE CIENCIAS**

**ACTUALIZACIÓN DEL MODELO DE  
CALIFICACIÓN PARA LA MEDICIÓN DEL  
RIESGO CREDITICIO DE LA CARTERA  
EMPRESARIAL**

**REPORTE DE TRABAJO  
PROFESIONAL**

QUE PARA OBTENER EL TÍTULO DE:

**ACTUARIO**

P R E S E N T A:

**JOSÉ CARLOS MONTES DE OCA HERNÁNDEZ**

**Tutor :**

**ACT. RODRIGO CELORIO ROMERO**

**2018**





Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



## Datos del Jurado

### 1. Datos del alumno

Montes de Oca  
Hernández  
José Carlos  
55 27 16 60 81  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Actuaría  
308602528

### 2. Datos del tutor

Act.  
Celorio  
Romero  
Rodrigo

### 3. Datos del Sinodal 1

M. en C.  
Benítez  
Sánchez  
Helen Patricia

### 4. Datos del Sinodal 2

M. en B.M.F.  
Moreno  
Cortés  
Aurora Cbeta



5. Datos del Sinodal 3

Act.  
Martínez  
Guerrero  
Yolanda

6. Datos del Sinodal 4

Act.  
Román  
Pedraza  
Julio César

7. Datos del trabajo escrito

Actualización del Modelo de Calificación para la Medición del Riesgo  
Crediticio de la Cartera Empresarial  
80 p.  
2018



## Índice

Introducción .....	6
Capítulo 1 Riesgo Crediticio .....	8
1.1 El Crédito .....	8
1.2 Antecedentes de los Modelos de Calificación Crediticia .....	9
1.3 Métodos para evaluar el riesgo crediticio .....	11
Modelo de Beaver .....	16
Modelo Z-score .....	16
Modelo Zeta .....	17
Modelos de Regresión .....	17
Matrices de Transición .....	17
Credit Scoring .....	19
1.4 Método Regulatorio .....	19
1.5 Métodos Usados por las Calificadoras .....	24
Capítulo 2 Metodologías .....	27
2.1 Análisis discriminante .....	27
2.2 Modelos de Regresión .....	30
2.2.1 Modelo de Probabilidad Lineal .....	30
2.2.2 Modelo Logit .....	31
2.2.3 Modelo Probit .....	33
2.3 Redes Neuronales .....	34
2.4 Árboles de Decisión .....	36
2.5 Análisis de Componentes Principales (ACP) .....	38
Capítulo 3 Construcción del Modelo de Calificación .....	48
3.1 Obtención y clasificación de la información de los clientes .....	48
3.2 Variables para utilizar .....	49
3.3 Reducción a las variables más discriminantes .....	52
3.4 Selección de variables .....	60
3.5 Obtención de la escala de calificación del modelo .....	60
Capítulo 4 Validación del Modelo .....	64
4.1 Comparación con modelos de agencias especializadas .....	64
4.2 Comparación del modelo contra el método regulatorio .....	65
Capítulo 5 Resultados .....	67



---

Capítulo 6 Conclusiones .....	74
Bibliografía.....	76
Anexos .....	78
Glosario.....	80



## Introducción

El presente trabajo profesional el cual lleva por nombre “*Actualización del Modelo de Calificación para la Medición del Riesgo Crediticio de la Cartera Empresarial*”, tiene por objeto describir las actividades llevadas a cabo con el objetivo de actualizar el modelo interno de calificación que utiliza una institución financiera en su estrategia para el otorgamiento, *pricing* y supervisión de los financiamientos.

Este proceso se llevó a cabo verificando la consistencia, integridad, precisión y validez de la información utilizada. Así como la veracidad de las fuentes para la correcta Administración del riesgo crediticio en la cartera empresarial.

Como se podrá observar a lo largo del presente trabajo, es crucial para una mejor administración del riesgo crediticio, que el modelo a ocupar a la hora de otorgar, dar una tasa, dar seguimiento y supervisar un crédito, se encuentre actualizado y validado, al mismo tiempo, que éste sea confiable y preciso para evitar el posible deterioro de la cartera crediticia, así como, la prevención de pérdidas en que puedan incurrir los acreditados de la institución.

El presente reporte estará integrado por seis capítulos, y se contempla cada uno conforme a lo siguiente:

El primer capítulo del informe estará concentrado en cuatro rubros principales: ¿qué es un crédito?, las principales formas de evaluar el riesgo crediticio, la regulación que rige en el país y los antecedentes de algunos modelos para la medición del riesgo de crédito.

En segunda instancia se ve la teoría de las metodologías más utilizadas para la medición del riesgo crediticio.



Posteriormente se construye el modelo y se verificará la significancia, la dependencia y la relación entre las variables a utilizar encontrando si son, o no, significativas para el modelo, observando cuales dan mejores resultados y mayor confiabilidad al modelo, para finalmente obtener la escala de calificación.

En el siguiente capítulo se valida el enfoque utilizado para la construcción del modelo y se realiza un comparativo teórico versus los modelos utilizados por la instancia reguladora en México, mismo que usualmente trabajan las calificadoras crediticias especializadas.

En el penúltimo capítulo se realizan escenarios sobre cuatro empresas con el modelo propuesto, comparando los resultados obtenidos con las calificaciones otorgadas por una agencia calificadora especializada y los resultados del modelo regulatorio de la Comisión Nacional Bancaria y de Valores.

Finalmente se aportan las conclusiones de llevar a cabo la actualización del modelo, los resultados obtenidos en comparación con otros y las mejoras a llevar a cabo para que el modelo siga funcionando como lo esperado.



## Capítulo 1 Riesgo Crediticio

### 1.1 El Crédito

Etimológicamente la palabra crédito tiene su origen del latín *creditus* el cual es un sustantivo del verbo *credere* (creer) que significa cosa confiada (Morales Castro y Morales Castro, 2014 :23), se utiliza la palabra crédito para referirnos a una cantidad de dinero que es puesta a nuestra disposición y que se debe devolver en un tiempo determinado.

En el caso de las instituciones financieras que permiten disponer de una cierta cantidad de dinero al prestatario, quien paga una compensación (comisión) por el préstamo del dinero, más una parte proporcional del riesgo en que la institución está incurriendo al confiar en él conocida como intereses de acuerdo con unas condiciones pactadas previamente. Cabe aclarar que por lo general en un crédito sólo se pagan intereses sobre el capital utilizado, el resto del dinero está a nuestra disposición, pero sin que por ello se tenga que pagar intereses. Llegado el plazo del vencimiento del crédito se puede volver a negociar su renovación o ampliación.

Los indicios del crédito se remontan a la Antigua Roma, ya que se han encontrado leyes o decretos que dejan constancia de penas corporales, enajenación de bienes o incluso hasta pena de muerte a los deudores que no cumplían con sus obligaciones. (Silva, 2016: 5)

Para que se diera el crédito como actualmente se conoce, fue necesario el surgimiento del dinero en forma de metales preciosos que tuvieran valor como bien en sí mismo, que fuera duradero, se pudiera dividir y transportar. En Europa regularmente fue el oro el elegido para usarse como moneda para grandes transacciones sin embargo no se encontraba seguro en algún lado. (Silva, 2016: 6)



Los orfebres quienes trabajaban este metal y contaban con cajas fuertes que lo resguardaran o con personas contratadas para protegerlo, fueron los que se encargaron de cuidarlo y resguardarlo, posteriormente la gente al darse cuenta de que no necesitaban cargar físicamente con el dinero al realizar un pago si no únicamente extender una carta al orfebre y que la persona pasar a recogerlo provocando que el sistema fuera más eficiente. De esta forma los orfebres se dieron cuenta que contaban con demasiado dinero que no era retirado de sus arcas a lo que decidieron volverlo productivo mediante el otorgamiento de créditos a personas que lo necesitaban, convirtiéndose así en los primeros banqueros. Sin embargo, se dieron cuenta de que se enfrentaban al problema de no saber cuándo sus clientes iban a recoger su dinero, teniendo el oro suficiente en esos casos por lo que prestaban una parte y la otra la reservaban. Otro problema era que no todos sus acreedores cumplían por lo que iniciaron así a estimar posibles pérdidas. (García Sánchez y Sánchez Barradas, 2005: 10)

Con el surgimiento de los créditos la cantidad de dinero disponible aumentó ya que es el dinero en circulación más la cantidad que se encuentra depositada, más el monto en créditos. Lo que provoca que exista una relación directa entre el dinero de una economía y la cantidad de reservas. (García Sánchez y Sánchez Barradas, 2005: 12)

## 1.2 Antecedentes de los Modelos de Calificación Crediticia

Para obtener la distribución crediticia de un portafolio lo primero a realizar es establecer un sistema de calificación en el cual cada acreditado obtenga una calificación con base en la probabilidad de que pueda caer en incumplimiento en sus obligaciones durante un cierto periodo de tiempo (generalmente el horizonte de tiempo es a un año). La calificación de riesgo crediticio es un indicador que toma en cuenta análisis cualitativos y cuantitativos del acreditado.

La calificación crediticia nació en el mercado de valores de EUA para diferenciar el riesgo de los títulos de valores marcando diferencia entre los sectores corporativos y municipales



para posteriormente expandirse a otros sectores y llega a Latinoamérica al inicio de la década de los 90's. (Venegas, 2011: 751,753)

El uso de las matrices de transición se ocupa para estimar el posible deterioro de un acreditado en cierto periodo de tiempo calculando la probabilidad de que el acreditado mejore o empeore su calificación. (Venegas, 2011: 785,786)

Se entiende que algún cambio en la calidad del acreditado y por lo tanto en su calificación crediticia genere una pérdida o una ganancia. Un ejemplo sería cierto acreditado, el cual tuvo una mejoría en su calificación otorgándole una sobretasa corporativa menor, ocasionando que el diferencial de riesgo o prima caiga, lo que le representa al acreditado una ganancia final. De estos cambios en la calidad de crédito, surgen las denominadas Pérdidas No Esperadas, que representan la principal medida de riesgo.

La probabilidad de incumplimiento definida con base en la CUB<sup>1</sup>, se entiende como la probabilidad expresada como porcentaje de que ocurra cualquiera o ambas de las siguientes circunstancias en relación a un deudor específico:

- a) El deudor se encuentra en situación de mora durante 90 días naturales o más respecto a cualquier obligación importante frente a la Institución. La Comisión podrá autorizar excepcionalmente el uso de un plazo diferente al de 90 días naturales o más para las Operaciones Sujetas a Riesgo de Crédito con las personas a que se refiere la fracción IV del Artículo 2 Bis 69 de las presentes disposiciones cuando, a su juicio, dicha definición de incumplimiento se ajuste mejor al método basado en calificaciones internas de que se trate.
- b) Se considere probable que el deudor no abone la totalidad de sus obligaciones crediticias frente a la Institución.

---

<sup>1</sup> Definición CXLII del Capítulo I del Título Primero, Disposiciones Generales



Generalmente un crédito puede sufrir ciertas alteraciones durante su ciclo de vida. Pudiéndose observar ciertos incumplimientos que únicamente son simples retrasos por pequeños desajustes de liquidez o por fallas operacionales de los acreditados. Por lo que un acreditado que cae en incumplimiento porque suspende sus pagos de pago es diferente de otro acreditado que deja de pagar por alguna falta temporal de liquidez y que posteriormente puede regularizar su situación. Por esto como lo dice la definición anterior el incumplimiento empieza cuando pasa cierto periodo de tiempo, en este caso 90 días. Conviene aclarar que el hecho de que un acreditado se vaya a incumplimiento no implica que el saldo a esa fecha vaya a ser la pérdida final de dicho crédito, pues este monto también se puede recuperar vía judicial o mediante reestructura de sus adeudos en cuyo caso no se registra ninguna pérdida aunque haya existido algún incumplimiento.

### 1.3 Métodos para evaluar el riesgo crediticio

La CNBV en su Circular Única de clasifica el Riesgo de Crédito como uno de los Riesgos Cuantificables, para los cuales es posible conformar bases estadísticas que permitan medir sus pérdidas potenciales, y dentro de estos como un Riesgo Discrecional, ya que es un riesgo resultante de la toma de una posición de riesgo, para finalmente definir al Riesgo de Crédito como la pérdida potencial por la falta de pago de un acreditado o contraparte en las operaciones que efectúan las Instituciones, incluyendo las garantías reales o personales que les otorguen, así como cualquier otro mecanismo de mitigación utilizado por las instituciones.

En los últimos años, el área de administración de riesgo ha tenido un gran desarrollo dado por la regulación a partir de las recientes crisis. Al principio, los modelos de riesgo se enfocaban en medir únicamente el riesgo de mercado que presentan los portafolios de inversión de las instituciones financieras. Las instituciones para reducir los requerimientos de capital que han sido impuestas por las autoridades regulatorias para evitar dichas crisis han provocado el enriquecimiento del área de administración de riesgos.



El desarrollo de metodologías que midan el riesgo de crédito ha sido menor en comparación con el riesgo de mercado, debido principalmente a las dificultades que presenta este tipo de riesgo, como la medición de los factores, la falta de información adecuada y el costo de elaborarla, han provocado obstáculos en la prevención y cuantificación del riesgo de crédito.

A pesar de esto se han propuesto distintas Metodologías para estimar la probabilidad de incumplimiento de los acreditantes, utilizándose por lo general Técnicas Estadísticas incluyendo la regresión lineal, regresión logística, análisis discriminante, redes neuronales, árboles de decisión, modelos probit y logit, análisis de matrices de transición y juicio experto.

El primero en proponer un modelo para medir el riesgo de crédito fue Beaver (1966), quien utilizó el análisis discriminante basándose en el comportamiento financiero de las empresas para predecir la probabilidad de quiebra en empresas corporativas.

El principal problema de las instituciones financieras es administrar su capital económico. Para ello, se busca un Modelo que combine riesgo y rendimiento.

La correcta gestión de los riesgos en que incurre una institución financiera es fundamental por lo que se necesita identificar, analizar, medir, vigilar, limitar, controlar, revelar y dar tratamiento a los riesgos a los que está expuesta dicha institución.

Dentro de la institución financiera la encargada de todas estas acciones es la Dirección de Administración de Riesgos, la cual mediante reportes periódicos mantiene informada a la Administración sobre el nivel de riesgo que está asumiendo la Institución, ya sea la Dirección General y/o el Consejo de Administración. Lo anterior, para tener argumentos en la toma de decisiones informadas y oportunas acerca de la prevención de pérdidas y la protección del capital de la institución.

Para una buena administración del riesgo de crédito se debe contar con modelos eficientes y actualizados que ayuden a una mejor gestión de la cartera crediticia, mayor competitividad, así mismo, se deben implementar métodos de valuación para asegurar que los modelos se encuentren bien calibrados.

A pesar de que se han desarrollado distintos métodos para eficientar y mejorar el otorgamiento y seguimiento de los créditos, no se tiene que perder de vista que el objetivo final de medir el riesgo de crédito es poder prever anticipadamente las pérdidas potenciales en las que se puede incurrir mediante el otorgamiento de créditos. Se resalta que el objetivo de todas las metodologías referentes al riesgo crediticio se basa en el cálculo de la probabilidad de incumplimiento de los acreditados.

Para tener una adecuada administración del riesgo se deben utilizar métodos, que en conjunto con unas buenas políticas de otorgamiento de crédito ayuden a detectar posibles incumplimientos antes de que ocurran y evitar asumir obligaciones con créditos que puedan generar problemas.

Con base en el Artículo 5 de la Circular Única de Bancos (CUB)<sup>2</sup> se pueden apreciar las siguientes etapas de un crédito:

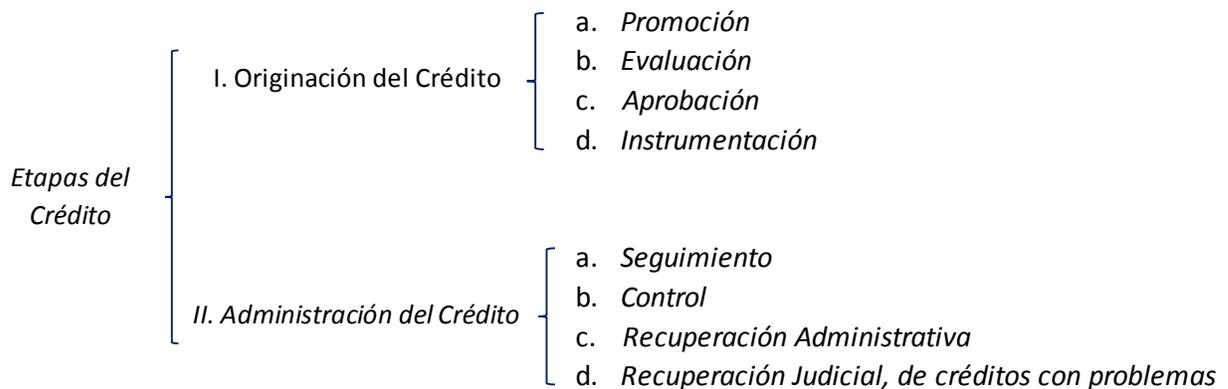


Imagen 1. Etapas de un crédito

<sup>2</sup> Disposiciones de Carácter General Aplicables a las Instituciones de Crédito



En la fase de evaluación se utilizan modelos cuantitativos para determinar la calificación de crédito que permite establecer el riesgo del crédito, así como los límites en el otorgamiento.

Fundamentalmente los modelos para estimar el riesgo de crédito se pueden dividir en dos tipos:

- **MODELOS TRADICIONALES:** Los cuales utilizan criterios subjetivos como el juicio o la experiencia del analista. Este procedimiento se reduce en las cinco Ces del solicitante (Conocer, Capacidad, Capital, Colateral y Condiciones Cíclicas).
- **MODELOS DE ENFOQUE MODERNO:** Buscan estimar tanto la Pérdida Esperada como la Pérdida No Esperada, ésta última es el nivel de capital mínimo que un banco debe tener para mantener su portafolio de créditos y a su vez se divide en dos:
  - a. En los modelos condicionales las metodologías buscan conocer las principales causas de incumplimiento basándose en las distintas variables financieras y macroeconómicas para realizar un modelo de correlaciones de causalidad entre ellas.
  - b. En los modelos no condicionales las metodologías buscan conocer las probabilidades de incumplimiento utilizando un cierto conjunto de variables que caracterizan a los individuos sujetos de crédito.

En los modelos tradicionales se han elaborado distintos procedimientos siendo el más conocido el de las cinco “Ces” del solicitante o modelo experto, donde la decisión de otorgar o no un crédito se basa en la experiencia del analista en la evaluación de los siguientes aspectos:

1. **Conocer al sujeto de crédito:** Se toma en cuenta su solvencia moral, reputación y disposición para cumplir sus obligaciones con terceros, es decir su historial de pagos. Un buen indicador de la reputación de la empresa es su antigüedad en el mercado.



2. **Capacidad de pago:** Con base en un análisis financiero que mida la habilidad para pagar, la cual se refleja en la volatilidad de las utilidades generadas, así como el flujo de efectivo que sea suficiente para pagar las obligaciones futuras.
3. **Capital:** Se observa el compromiso de los accionistas en el capital total de la empresa, así como la capacidad de endeudamiento. Si tienen un alto nivel de apalancamiento la probabilidad de quiebra aumenta.
4. **Colateral:** Son las garantías que otorga el deudor. Para en el caso de incumplimiento de pago deberán ser suficientes para que el banco recupere la pérdida en su inversión. A mayor valor de las garantías la exposición al riesgo de crédito es menor.
5. **Ciclo económico:** Es un elemento importante en la determinación de la exposición crediticia, sobre todo en aquellos sectores económicos que dependen de él. (De Lara, 2008: 163)

Se aclara que comúnmente el analista utiliza estos aspectos, más como una guía y puede tomar en cuenta más elementos para emitir un análisis más completo.

Se notan varios problemas en este modelo, como que la decisión del analista varíe en función de la información recabada convirtiendo en algo difícil la comparación de los rangos a tomar en cuenta en el análisis.

Por lo tanto, se suele complementar este modelo con metodologías del enfoque moderno ya que en grandes instituciones se busca conseguir que el análisis sea más homogéneo.

Los principales modelos para estimar el riesgo de crédito y la probabilidad de incumplimiento bajo el enfoque moderno, los cuales utilizan los modelos estadísticos descritos más adelante.



### **Modelo de Beaver**

Éste fue un modelo discriminatorio simple en el que se compararon las medias de 30 razones financieras de 158 empresas (la mitad de ellas en bancarrota) de 38 sectores económicos obteniendo que diversas razones podían predecir con un 90% de nivel de confianza o más la bancarrota con un año de anticipación. El Ingreso Neto sobre la Deuda Total tuvo un 92% de precisión, seguido del Ingreso Neto sobre las Ventas con un 91% y después de estos con un 90% quedaron el Ingreso Neto sobre Patrimonio Neto, Flujo de Caja sobre Deuda Total y Flujo de Caja sobre Total de Activos. (Restrepo y Orozco, 2015: 12)

### **Modelo Z-score**

Es un modelo que analizó la información financiera de compañías manufactureras en un análisis discriminante múltiple diseñado por Edward I. Altman, en el cual tomó datos de 66 empresas (la mitad de ellas que se fueron a bancarrota) con un horizonte de tiempo de 19 años (1946-1965) tomando como base 22 razones financieras, pudiendo predecir un 72% de las empresas que se fueron a quiebra con dos años de anticipación. El modelo dio forma a una función  $z$  que toma en cuenta 5 variables asignándoles un coeficiente a cada una de ellas y con base al resultado obtenido las divide en posible quiebra, indefinida y en buen estado. (Amat, 2014)

Desde los trabajos publicados por Altman en los años setenta hasta mediados de los años noventa en general no se implementaron nuevos modelos, los modelos ocupados en esta época solo contenían pequeñas variaciones del modelo Z-score de Altman adaptados a las distintas necesidades como el modelo Credit –Man o el de Edminster que comparan a las empresas con la media de su sector o como el de Conan y Holder enfocado a Pymes francesas.

Este modelo presenta algunas inconveniencias ya que al utilizar razones financieras se provoca un efecto de subestimación en el tamaño de las estadísticas no pudiendo identificar



datos relevantes en el otorgamiento de un crédito. Así mismo en este modelo se ocuparon empresas grandes por lo que su aplicación tanto para compañías pequeñas y medianas no resulta muy satisfactoria.

### **Modelo Zeta**

En 1977, Altman, Haldeman y Narayanan hacen algunas modificaciones al modelo Z-Score, entre ellas, en el análisis utilizan empresas de todos los tamaños, empresas de otros sectores y no solo del manufacturero y técnicas de estadística más recientes. Resultando en un modelo superior al original que permite predecir la bancarrota de una empresa con un de anticipación utilizando un nivel de confianza del 96%. (Amat, 2014)

El problema con esta clase de modelos es que se asume que la información reportada en los estados financieros asume una distribución normal y que la matriz asociada de varianzas y covarianzas de empresas en estado de quiebra y los que no lo están son iguales lo que en la mayoría de los casos no se cumple.

### **Modelos de Regresión**

En esta clase de modelos se busca una ecuación que permita obtener la probabilidad de que un determinado evento ocurra, en función de las características que defina el analista que la realiza. Estos modelos son conocidos normalmente como modelos Probit y Logit. Estos se basan en que existe una relación lineal entre la variable dependiente y las independientes, permitiéndonos identificar y cuantificar la probabilidad de incumplimiento para así poder estimarla.

### **Matrices de Transición**

Las matrices de transición se comenzaron a utilizar a partir de 1997 para medir el riesgo de crédito apareciendo por vez primera en el modelo *CreditMetrics* desarrollado por un grupo



de instituciones financieras encabezadas por JP Morgan. Convirtiéndose en uno de los modelos más utilizados debido principalmente a su fácil implementación. (De Lara, 2008: 183)

Mediante el uso de matrices de transición se calcula la probabilidad de que un crédito mantenga, mejore o empeore la calificación otorgada en un cierto periodo de tiempo. Llamándole a esta probabilidad de migración en la calidad de un crédito.

Se hace uso de los procesos estocásticos estimando la probabilidad de pasar de un estado (i) el cual se refiere a la calificación designada del crédito a un cierto periodo de tiempo (t), a un estado (j) en el siguiente periodo de tiempo (t+1).

Se construye una matriz de transición que tenga i filas y j columnas, tomando en cuenta que debe satisfacer las siguientes restricciones:

- Los elementos deberán ser positivos
- La suma de cada fila es igual a 1

La utilidad de este modelo en la estimación del riesgo crediticio es la ayuda que otorga para poder prever el deterioro de los créditos.

Los dos modelos internos mayormente reconocidos y utilizados son *CreditMetrics* y *Creditrisk* publicados en 1997 por los bancos JP Morgan y Credit Suisse, respectivamente. Ambos modelos adoptaron la teoría ya existente de portafolios enfocándola al riesgo crediticio, construyen carteras de crédito diversificando sus créditos para obtener una medida de Valor en riesgo (VaR); sin embargo, las diferencias entre los dos tipos de riesgo complican su completa asimilación. (De Lara, 2008: 190)



### ***Credit Scoring***

“Los métodos o modelos de *credit scoring*, a veces denominados *score-cards* o *classifiers*, son algoritmos que de manera automática evalúan el riesgo de crédito de un solicitante de financiamiento o de alguien que ya es cliente de la [institución financiera].” (Gutiérrez, 2007: 3)

Derivado del gran número de financiamientos que otorgan las instituciones financieras, este método es de los más ocupados ya que puede otorgar una clasificación rápida reduciendo el costo de análisis.

En estos modelos se estima una calificación al crédito en base al cumplimiento de las obligaciones crediticias pasadas del solicitante, así como a su información financiera y personal. Para poder llevar a cabo esta asignación previamente utilizando métodos estadísticos o matemáticos se realizó la evaluación de las variables que se ocuparan, así como del peso individual en el modelo de *scoring*. Esto para permitir modelar el riesgo de incumplimiento de operaciones futuras.

#### **1.4 Método Regulatorio**

En 1988, se lleva a cabo el Primer Acuerdo de Basilea, conocido comúnmente como Basilea I, el cual emite recomendaciones sobre la regulación bancaria. El Comité de Basilea está formado por representantes de los Bancos Centrales y de las Autoridades Regulatorias del Grupo de los Diez más Luxemburgo y España. En 2009, se aceptaron como miembros los países del G-20 que aún no tenían representación, así como las plazas bancarias de Hong Kong y Singapur. Cabe aclarar que el Comité no tiene autoridad para imponer recomendaciones, si bien la mayor parte de los países, así como algunos otros que no forman parte de este tienden a implementar sus políticas. Esto significa que las recomendaciones son aplicadas a través de leyes y regulaciones nacionales. (*Bank for International Settlements, Historia*)



El objetivo de este Primer Acuerdo era determinar el capital mínimo económico sujeto a riesgo de crédito, en el cual se estableció que el requerimiento mínimo de capital fuera del 8%, obligando así a una recapitalización de los bancos.

Este capital mínimo se implementó en México en enero de 1993, observándose que este criterio omite tres aspectos imprescindibles para la correcta medición del riesgo de crédito. No toma en cuenta la calidad crediticia de los acreditados; no hace diferencia en el plazo restante del crédito respecta al monto y no toma en cuenta la diversificación del portafolio.

En el año 2000 se complementó la regulación incorporando también el riesgo de mercado, el cual fue sumado al riesgo de crédito para cumplir con el requerimiento del 8% del capital.

Los tres aspectos anteriormente mencionados, que no fueron tomados en cuenta en Basilea I para la medición del riesgo de crédito provocaron que las instituciones financieras con el fin de mejorar la medición su capital económico y poder administrar óptimamente su capital desarrollaran modelos internos para calificar sus créditos, los cuales fueron los precursores de la clasificación que se incorporó en Basilea II.

En el Segundo Acuerdo de Basilea acontecido en 2004 (Basilea II) se incorporó el requerimiento mínimo de capital, el Riesgo Operacional, así como, la incorporación de la Supervisión y Revelación del Mercado, lo que incluye que los supervisores regulatorios cooperen entre sí, que las instituciones financieras revelen y entreguen reportes a la autoridad de su composición de capital y la exposición al riesgo. Y en el cual ya aparece la clasificación de los créditos como el IRB (Internal Ratings-Based) para que los bancos puedan utilizar su propia clasificación de crédito cumpliendo ciertas condiciones bajo la supervisión de la autoridad regulatoria. Estas medidas fueron implementadas en México en el año 2007 y posteriormente la clasificación fue modificada en el año 2010 bajo la metodología de pérdidas esperadas. (Basel Comitee on Banking Supervision, 2005: 9)



En el año 2010 surgió el Tercer Acuerdo de Basilea (Basilea III) derivado de la acumulación de fallos en el sistema financiero, los cuales quedaron patentes en la crisis financiera del 2008. En este nuevo acuerdo se introdujo un Buffer anti cíclico de capital de 2.5%, elevando el requerimiento mínimo de capital de los tres tipos de riesgo definidos en los acuerdos pasados a 10.5%, se incorporó la razón de apalancamiento y se adicionaron requerimientos de liquidez.

En México la institución encargada de emitir y verificar el cumplimiento de la regulación en el sistema financiero es la Comisión Nacional Bancaria y de Valores.

La regulación que rige actualmente en México para el cálculo de la calificación de cartera crediticia y la generación de sus reservas preventivas de riesgo de crédito para cada crédito data del año 2014 la cual, siguiendo el enfoque de Basilea, divide a las empresas en dos tipos, de acuerdo a su tamaño de ingresos netos o ventas netas en empresas y grandes empresas, lo cual para el caso de México se estableció el límite para dividir las en 14 millones de UDIS<sup>3</sup>, cambiando los aspectos y las ponderaciones que se encargan de asignar la calificación de cada uno de acuerdo a su tamaño.

Este enfoque, así mismo, está basado en los últimos lineamientos establecidos por el Comité de Basilea que asigna una calificación a cada acreditado de la institución basado en los distintos niveles de riesgo de acuerdo a la pérdida esperada que se le calcula a cada crédito.

El cálculo del puntaje crediticio que obtiene cada crédito para calcular su reserva regulatoria se detalla en la Circular Única de Bancos en sus anexos 21 y 22.

---

<sup>3</sup> UDIS. Unidades de inversión. Son los factores que se aplican a la adquisición de créditos hipotecarios y se publican diariamente por el Banco de México en el Diario Oficial de la Federación. En lo que va del año (hasta el 30 de agosto) el promedio de la UDIS se ubica en 6.01 MN con un máximo de 6.07 y un mínimo de 5.94. Usando el promedio 14 millones de UDIS equivalen a 84'149,715.19 MN



Con la información proporcionada en el anexo 21 se puede determinar el puntaje crediticio de los créditos concedidos a personas morales (distintas a entidades federativas, municipios y entidades financieras) y a personas físicas con actividad empresarial, cuyos ingresos netos o ventas netas<sup>4</sup> anuales sean menores al equivalente en moneda nacional a 14 millones de UDIS. Dentro de este apartado se dividen a los acreditados en dos categorías: sin atraso en pagos y con atraso en pagos de acuerdo a la información existente en las Sociedades de Información Crediticia<sup>5</sup>. En caso de no contar con esta se podrá utilizar información interna de la institución de los últimos 12 meses. Si esta tampoco existe se considerará como Acreditado con Atraso. (Anexo 21 CUB, 2018: 1)

El puntaje crediticio total de este anexo únicamente se basa en un puntaje cuantitativo que se calcula mediante la suma de los 3 factores de riesgo que lo componen, los cuales son:

- Factor de riesgo experiencia de pago. Compuesto por 12 indicadores.
- Factor de riesgo experiencia de pago INFONAVIT. Compuesto por 2 indicadores.
- Factor de riesgo características propias de la empresa. Compuesto por 3 indicadores.

Con la información proporcionada en el anexo 22 se puede determinar el puntaje crediticio de los créditos concedidos a personas morales (distintas a entidades federativas, municipios y entidades financieras) y personas físicas con actividad empresarial, con ingresos netos o ventas netas anuales mayores o iguales al equivalente en moneda nacional a 14 millones de UDIS. Adicionalmente en este anexo se encuentra una subdivisión extra para calcular el puntaje total dividiéndolos en tres tamaños: el primero para pequeños corporativos que van de 14 millones de UDIS a 54 millones de UDIS en ventas netas, corporativos de 54 millones de UDIS a 216 millones de UDIS en ventas netas y en grandes corporativos que son los que tienen ventas netas mayores a 216 millones a UDIS. (Anexo 22 CUB, 2018: 1)

---

<sup>4</sup> Ingresos que genera una entidad por la venta de inventarios, prestación de servicios, o por cualquier otro concepto que se deriva de las actividades primarias que representan la principal fuente de ingresos de la propia entidad, menos descuentos, bonificaciones y devoluciones (NIF B-3)

<sup>5</sup> En el caso de México operan dos: el Buró de Crédito y el Círculo de Crédito. Ambas son empresas privadas que reciben información de quienes otorgan préstamos y la transforman en historiales crediticios.



El puntaje de riesgo cuantitativo en este anexo se obtiene mediante la suma de 3 factores de riesgo, que son:

- Factor de riesgo experiencia de pago. Compuesto por 4 indicadores.
- Factor de riesgo experiencia de pago INFONAVIT. Compuesto por 3 indicadores.
- Factor de riesgo características propias de la empresa. Compuesto por 3 indicadores.

El puntaje de riesgo cualitativo en este anexo se obtiene mediante la suma de 5 factores de riesgo, que son:

- Riesgo país y de la industria. Compuesto por 2 indicadores.
- Posicionamiento del mercado. Compuesto por 2 indicadores.
- Transparencia y estándares. Compuesto por 2 indicadores.
- Gobierno corporativo. Compuesto por 3 indicadores.
- Competencia de la administración. Compuesto por 2 indicadores.

En este anexo el puntaje total es de acuerdo al tamaño de la empresa ya mencionado anteriormente. Para Pequeños Corporativos la calificación total es únicamente el puntaje que arroje el riesgo cuantitativo, para Corporativos la calificación total la da en 75% el riesgo cuantitativo y el 25% restante lo otorga el riesgo cualitativo y finalmente para obtener la calificación de los Grandes Corporativos el 65% lo da el riesgo cuantitativo y el 35% el riesgo cualitativo.

El problema de usar esta calificación a la hora del otorgamiento y *pricing* de los créditos es que no toma en cuenta la situación financiera de la empresa, por ejemplo a las empresas calificadas con base al anexo 21, de los 17 indicadores que ocupa para asignar la calificación regulatoria, únicamente 3 corresponden a características propias de la empresa, pero de estas ninguna refleja absolutamente nada respecto a la situación financiera que atraviesa la empresa, así como de su capacidad de pago por lo que en caso de que dicha empresa cayera en incumplimiento y hubiera cumplido puntualmente con todos sus pagos anteriores



no hubiera existido forma de que el banco se haya aprovisionado lo suficiente o hubiese tomado en cuenta otras medidas, lo que se traduce en una pérdida de capital no esperada.

Ahora veamos los ejemplos de los acreditados que por su tamaño de ventas les aplique el anexo 22 para otorgarles su calificación regulatoria, en el caso de los pequeños corporativos la calificación toma en cuenta 10 indicadores que los califican, 3 de ellos son razones financieras, aunque ninguna de ellas aporta información sobre su capacidad de pago, en el caso de las empresas calificadas como corporativos, para obtener su calificación ya se toman en cuenta aspectos cualitativos los cuales representan 11 indicadores más que ponderan el 25% de la calificación final, pero únicamente uno de ellos se refiere a capacidad de pago dando un peso promedio de 2.9% en la calificación final. Finalmente, para el caso de los grandes corporativos los 11 indicadores del aspecto cualitativo ponderan el 35% de la calificación final, de la cual el indicador que mide la capacidad de pago tiene un peso promedio de 4.1% en la calificación final.

Como se puede observar, el indicador que tiene mayor peso a la hora de otorgar la calificación regulatoria es la **experiencia de pago** calificando a la mayor parte de las empresas con base en esto y no tanto al desempeño como tal de la empresa lo que nos lleva a la necesidad de tener una metodología que si tome en cuenta aspectos financieramente relevantes de la empresa que indiquen su desempeño, así como variables que también se consideren relevantes para la toma de decisiones del otorgamiento de los créditos.

### 1.5 Métodos Usados por las Calificadoras

Las calificaciones que reportan las agencias calificadoras comúnmente son llamadas riesgo de emisor, la cual es una evaluación de la vulnerabilidad de algún emisor al incumplimiento de sus obligaciones financieras, pretendiendo con esta calificación poder comparar empresas de distintos países, así como de distintos grupos de la industria. Las calificaciones pueden ser de largo y/o corto plazo. (Fitch Ratings, 2017: 1)



Dichas calificaciones reflejaran factores cualitativos y cuantitativos comprendiendo riesgos de negocio y financieros.

Los principales factores para evaluar son:

- Perfil de riesgo del sector
- Ubicación
- Estrategia de administración/gobierno corporativo
- Estructura del grupo
- Perfil de negocio
- Perfil Financiero
  1. Flujo de efectivo y rentabilidad
  2. Estructura de capital
  3. Flexibilidad financiera

Se califican las proyecciones en un horizonte de tiempo de 3 a 5 años con una historia de 3 años.

Tomando como base el perfil del riesgo del sector, la ubicación, así como, las características operacionales y financieras se le determina un grupo apropiado, el cual limita la calificación a un nivel específico, el cual ayuda a que el criterio sea global.

La calificación obtenida se sensibiliza utilizando el criterio experto de los analistas tomando en cuenta el riesgo inherente de la industria, del negocio, de la ubicación, así como el financiero, para que dicha calificación sea prospectiva y esté enfocada a resultados futuros.

En el Perfil de Riesgo del Sector se utilizan una serie de “Navegadores de Calificación” llamados “Navigators”, los cuales reúnen elementos claves observados o previstos de las industrias, el cual aunque no establece un máximo o mínimo de calificación fija da una idea de cómo se desenvuelve el sector. (Fitch Ratings, 2017: 2)



En la ubicación se evalúa el entorno operativo, el cual toma en cuenta la ubicación de los ingresos y activos, así como la gobernanza en dicha ubicación, así como su riesgo de transferencia y convertibilidad, el cual se basa en la calificación soberana de dicha ubicación y los techos país. (Fitch Ratings, 2017: 3)

En estrategia de la administración se observa la capacidad en base al desempeño financiero, evaluando los objetivos corporativos, su financiamiento y expansión, para el gobierno corporativo se identifica la estructura del gobierno, la organizacional y la transparencia financiera, enfocándose en controles internos, políticas sólida, consejo directivo independiente y efectivo e integridad de los procesos de contabilidad y auditoría. . (Fitch Ratings, 2017: 4)

En estructura del grupo se califica a la tenedora o “holding” del grupo en caso de tenerla, evaluándose generalmente los compromisos o garantías existentes con sus subsidiarias o con empresas relacionadas; en caso de no evaluarse a una “holding”, se toma en cuenta la sustentabilidad y previsibilidad de los flujos de ingresos, así como su relación con la matriz del grupo y su contribución al perfil del grupo.

En Perfil del negocio se observa su posición en mercado clave, nivel de liderazgo en oferta de productos y capacidad de influenciar precios, así como varios factores claves distinto adaptados a los parámetros de cada sector. (Fitch Ratings, 2017: 5)

En el Perfil financiero en base a las cifras históricas se realizan proyecciones internas en el corto plazo, observándose principalmente el flujo de efectivo de las utilidades, la cobertura y el apalancamiento. (Fitch Ratings, 2017: 6)



## Capítulo 2 Metodologías

Varios autores han propuesto distintos métodos o modelos para medir el riesgo de crédito usando distintos enfoques y adaptando metodologías usadas en otros tipos de riesgos, basándose principalmente en la información financiera, se hablará y se darán las bases teóricas de algunos de ellos.

### 2.1 Análisis discriminante

Entre los modelos paramétricos, se encuentra: el análisis discriminante, el cual es una técnica estadística multivariada la cual analiza las diferencias existentes entre los grupos de objetos respecto a un conjunto de variables para obtener una función que pueda clasificar a un nuevo objeto a partir del conocimiento de ciertas variables. (Marín, s.f., Tema 6:1)

Mediante este análisis se buscan relaciones lineales entre las variables que puedan discriminar a los grupos dados permitiendo que se pueda asignar un nuevo objeto cuya clasificación se ignora a uno de los grupos.

Para llevar a cabo este análisis existen algunos supuestos a tomar en cuenta:

- i. Se tiene una variable categórica y el resto de las variables son independientes respecto de ella.
- ii. Se necesitan al menos dos grupos, y para cada grupo se necesitan mínimo dos casos.
- iii. El número de variables discriminantes debe ser menor que el número de objetos menos dos, es decir,  $(x_1, x_2, \dots, x_p)$  donde  $p < (n-2)$  siendo  $n$  el número de objetos.
- iv. Ninguna variable discriminante puede ser combinación lineal de otra.
- v. El número máximo de funciones discriminantes es igual al mínimo entre el número de variables y el número de grupos menos uno con  $q$  grupos,  $(q-1)$  funciones discriminantes.



- vi. Las matrices de covarianzas dentro de cada grupo deben de ser aproximadamente iguales.
- vii. Las variables continuas deben de seguir una distribución normal multivariante.

Teniendo q grupos a los que se les asigna a una serie de objetos y 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  $(y_1, y_2, \dots, y_m)$  de modo que sean funciones lineales de  $(x_1, x_2, \dots, x_p)$ :

$$\begin{aligned}
 y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p + a_{10} \\
 &\quad \dots \\
 y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mp}x_p + a_{m0}
 \end{aligned}
 \quad m = \min(q - 1, p)$$

tales que discriminen lo máximo posible a los q grupos.

Las combinaciones lineales de las p variables deben maximizar la varianza entre los grupos y minimizar la varianza dentro de los grupos.

La variabilidad total de la muestra se puede descomponer en variabilidad dentro de los grupos y entre los grupos. Para ello, se parte:

$$\text{Cov}(x_j, x_{j'}) = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j) (x_{ij'} - \bar{x}_{j'})$$

Se puede considerar la media de la variable  $x_j$  en cada uno de los grupos  $I_1, I_2, \dots, I_q$ , es decir:

$$\bar{x}_{kj} = \frac{1}{n_k} \sum_{i \in I_k} x_{ij} \quad \text{para } k=1, \dots, q$$

De esta forma, la media total de la variable  $x_j$  se puede expresar como función de las medias dentro de cada grupo:

$$\sum_{i \in I_k} x_{ij} = n_k \bar{x}_{kj}$$



Donde: 
$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} = \frac{1}{n} \sum_{k=1}^q \sum_{i \in I_k} x_{ij} = \frac{1}{n} \sum_{k=1}^q n_k \bar{x}_{kj} = \sum_{k=1}^q \frac{n_k}{n} \bar{x}_{kj}$$

Así, 
$$\text{Cov}(x_j, x_{j'}) = \frac{1}{n} \sum_{k=1}^q \sum_{i \in I_k} (x_{ij} - \bar{x}_j) (x_{ij'} - \bar{x}_{j'})$$

Poniendo en cada uno de los términos;

$$\begin{aligned} (x_{ij} - \bar{x}_j) &= (x_{ij} - \bar{x}_{kj}) + (\bar{x}_{kj} - \bar{x}_j) \\ (x_{ij'} - \bar{x}_{j'}) &= (x_{ij'} - \bar{x}_{kj'}) + (\bar{x}_{kj'} - \bar{x}_{j'}) \end{aligned}$$

Se obtiene:

$$\begin{aligned} \text{Cov}(x_j, x_{j'}) &= \frac{1}{n} \sum_{k=1}^q \sum_{i \in I_k} (x_{ij} - \bar{x}_j) (x_{ij'} - \bar{x}_{j'}) + \sum_{k=1}^q \frac{n_k}{n} (\bar{x}_{kj} - \bar{x}_j) (\bar{x}_{kj'} - \bar{x}_{j'}) \\ &= d(x_j, x_{j'}) + e(x_j, x_{j'}) \end{aligned}$$

La covarianza total es igual a la covarianza dentro de los grupos más la covarianza entre grupos. Si denominamos a  $t(x_j, x_{j'})$  como la varianza total entre las variables  $x_j$  y  $x_{j'}$  sin hacer una distinción de grupos lo anterior se puede expresar como:

$$t(x_j, x_{j'}) = d(x_j, x_{j'}) + e(x_j, x_{j'})$$

La idea básica del Análisis Discriminante consiste en extraer a partir de  $(x_1, x_2, \dots, x_p)$  variables observadas en  $k$  grupos,  $m$  funciones  $(y_1, y_2, \dots, y_m)$  de forma que:

$$Y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p + a_{i0}$$

donde  $m = \min(q - 1, p)$ , tales que  $\text{corr}(y_i, y_j)$  para todo  $i \neq j$

Si las variables  $x_1, x_2, \dots, x_p$  están tipificadas, implica que las funciones



$$Y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p \quad \text{para } i = 1, \dots, m$$

se denominan *discriminantes canónicas*.

Las funciones  $y_1, y_2, \dots, y_m$  se extraen de modo que:

- $y_1$  sea la combinación lineal de  $x_1, x_2, \dots, x_p$  proporcionando la mayor discriminación posible entre los grupos.
- $y_2$  sea la combinación lineal de  $x_1, x_2, \dots, x_p$  que proporciona la mayor discriminación posible entre los grupos, después de  $y_1$ , tal que  $\text{corr}(y_1, y_2) = 0$

En general,  $y_i$  es la combinación lineal de  $(x_1, x_2, \dots, x_p)$  que proporciona la mayor discriminación posible entre los grupos, después de  $y_{i-1}$  tal que  $\text{corr}(y_i, y_j) = 0$  para  $j = 1, \dots, (i - 1)$

## 2.2 Modelos de Regresión

En este tipo de modelos la variable dependiente es binaria pudiendo tomar únicamente valor de cero o uno. Generalmente se dividen en tres formas generales:

- Modelo de Probabilidad lineal.
- Modelo Logit, donde la función de distribución que siguen los errores es logística.
- Modelo Probit, cuando las perturbaciones se asumen con distribución normal.

### 2.2.1 Modelo de Probabilidad Lineal

En este tipo de modelo la variable dependiente es dicotómica y es función de las variables  $X_i$ . (Silva, 2016:28)

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$



La distribución en este tipo de modelo muestra una nube de puntos dividiendo a las observaciones en dos subgrupos. Los casos donde ocurrió el hecho  $Y_i=1$  y los casos donde no  $Y_i=0$ . Lo que provoca que el Coeficiente de Determinación  $R^2$  no sea muy ilustrativo ya que no aparecerán todos los datos en la recta de regresión.

Este modelo presenta algunas limitaciones al ser estimado por mínimos cuadrados como valores limitados por la respuesta dicotómica, la normalidad de las perturbaciones no asegura una distribución normal, hay estimaciones no acotadas y puede existir heterocedasticidad lo que restringe la bondad en el modelo.

### 2.2.2 Modelo Logit

Este modelo se basa en la función logística y se da de la siguiente forma: (Silva, 2016:31)

$$E(y) = \frac{e^{x'\beta}}{1 + e^{x'\beta}} + \varepsilon_i$$

Donde la  $x$  es el vector de variables independientes y  $\beta$  es el vector de parámetros, y también puede expresarse como:

$$\pi_1 = \frac{1}{1 + e^{-x'\beta}} + \varepsilon_i$$

Una equivalencia es:

$$1 - \pi_1 = \frac{1}{1 + e^{x'\beta}} + \varepsilon_i$$

Y de ambas podemos obtener la siguiente transformación:

$$\frac{\pi_1}{1 - \pi_1} = \frac{1 + e^{x'\beta}}{1 + e^{-x'\beta}} + \varepsilon_i = e^{x'\beta} + \varepsilon_i$$

Aplicando el logaritmo natural obtenemos:

$$\ln \frac{\pi_1}{1 - \pi_1} = x'\beta + \varepsilon_i$$

Con lo cual se tiene que el logaritmo de la razón de probabilidades es lineal en las variables y los parámetros.

La estimación de los parámetros de este modelo es vía máxima verosimilitud.



La forma general del modelo logit se puede expresar como:

$$y_i = E(y_i) + \varepsilon_i$$

Donde las observaciones  $Y_i$  son variables aleatorias independientes Bernoulli, con valores esperados:

$$E(y_i) = \pi_i = \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}}$$

Como cada observación sigue una distribución Bernoulli, su distribución será:

$$f_i(y_i) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i}, i = 1, 2, 3, \dots, n$$

Como las observaciones son independientes, la función de verosimilitud será:

$$L(y_1, y_2, \dots, y_n, \beta) = \prod_{i=1}^n f_i(y_i) = \prod_{i=1}^n \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$$

Al aplicarle logaritmo a la función de verosimilitud:

$$\ln L(y_1, y_2, \dots, y_n, \beta) = \ln \prod_{i=1}^n f_i(y_i) = \sum_{i=1}^n \left[ y_i \ln \left( \frac{\pi_i}{1 - \pi_i} \right) \right] + \sum_{i=1}^n \ln(1 - \pi_i)$$

Como:

$$1 - \pi_i = \frac{1}{1 + e^{x_i'\beta}} \text{ y } \ln \left( \frac{\pi_i}{1 - \pi_i} \right) = x_i'\beta$$

El logaritmo de verosimilitud se puede expresar para el modelo de regresión logística:

$$\ln L(y, \beta) = \sum_{i=1}^n y_i x_i'\beta - \sum_{i=1}^n \ln(1 + e^{x_i'\beta})$$

En muchas aplicaciones del modelo se dispone de información repetidas para cada uno de los valores de las variables. Sea  $Y_i$  la cantidad de 1 observados para la  $i$ -ésima observación y  $n_i$  la cantidad de ensayos en cada observación, entonces el logaritmo de la verosimilitud se puede presentar:

$$\ln L(y, \beta) = \sum_{i=1}^n y_i \pi_i + \sum_{i=1}^n n_i \ln(1 - \pi_i) - \sum_{i=1}^n y_i \ln(1 - \pi_i)$$

Los estimadores de máxima verosimilitud se pueden obtener mediante un algoritmo de mínimos cuadrados iterativamente re ponderados.



Si  $\hat{\beta}$  es el estimador obtenido, mediante el método iterativo y siendo ciertas las hipótesis del modelo, se puede demostrar que en forma asintótica:

$$E(\hat{\beta}) = \beta \text{ y } V(\hat{\beta}) = (X'V^{-1}X)^{-1}$$

El valor estimado del predictor lineal es  $\hat{\eta}_i = X_i\hat{\beta}$ , y el valor esperado del modelo de regresión logístico, se suele expresar:

$$\hat{y}_i = \hat{\pi}_i = \frac{e^{\hat{\eta}_i}}{1 + e^{\hat{\eta}_i}} = \frac{e^{x_i'\hat{\beta}}}{1 + e^{x_i'\hat{\beta}}} = \frac{1}{1 + e^{-x_i'\hat{\beta}}}$$

### 2.2.3 Modelo Probit

Este modelo se basa de la función de distribución de la normal tipificada, y el modelo se expresa de la siguiente forma: (Restrepo y Orozco, 2015:33)

$$\phi(z) = P(Z \geq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{s^2}{2}} ds + \varepsilon_i$$

Donde  $Z \sim N(0,1)$  y es tal que, dados los valores  $X_2, \dots, X_k$  de las variables independientes, se verifica que:

$$P(Y = 1 | x_2, \dots, x_k) = \phi(z_i)$$

Con  $z_i = \beta_1 + \beta_2 x_{2i} + \dots + \beta_k x_{ki}$  tal que:

$$Y = \begin{cases} 1 & \text{si } z_i > 0 \\ 0 & \text{si } z_i < 0 \end{cases}$$

Dadas las similitudes que presentan ambas curvas (normal tipificada y logística) los resultados estimados por ambos modelos discrepan únicamente en la rapidez con las que la curvas se aproximan a sus valores extremos, alcanzando la normal más rápidamente sus valores extremos.



## 2.3 Redes Neuronales

Una técnica que ha estado teniendo auge en los últimos años es el análisis de datos mediante el modelo de redes neuronales artificiales, el cual está basado en la función del cerebro humano, las representaciones matemáticas de varios de estos tipos de redes neuronales son muy parecidas a las del análisis estadístico tradicional.

Un ejemplo de la equivalencia de estos modelos es el de perceptrón simple con función lineal para una regresión lineal múltiple, un perceptrón simple con función logística para la regresión logística, un perceptrón simple con función umbral para la función discriminante lineal y un perceptrón multicapa auto asociativo para el análisis de componentes principales.

Estos modelos, a excepción del último utilizan perceptrones simples como base y puede ser visto como un Modelo Lineal Generalizado ya que la función de la activación de la neurona de salida en un perceptrón se define de la siguiente manera: (Montaño, 2002: 42,53)

$$Y = f(X, W)$$

El valor de Y se obtiene de una función de activación sobre una combinación lineal de coeficientes W y variables explicativas X.

Para obtener el perceptrón se intenta minimizar la suma del error cuadrático:

$$E = \sum_{p=1}^p \frac{1}{2} \sum_{k=1}^n (d_{pk} - y_{pk})^2$$

Con P indicando el elemento de distintos patrones posibles y n es el número de neuronas de salida,  $d_{pk}$  es la salida que se desea para la neurona k para el empleado y  $y_{pk}$  es la salida obtenida por el modelo.



Para intentar minimizar la función de error se ocupa la siguiente función:

$$E = - \sum_{p=1}^p \sum_{k=1}^n (d_{pk} \log y_{pk} + (1 - d_{pk}) \log(1 - y_{pk}))$$

Aunque en general en este tipo de modelos no se toma en cuenta el tipo de distribución de los errores ya que la obtención de los parámetros es mediante el criterio de optimización.

Una enorme diferencia entre este tipo de modelos y los de la estadística tradicional es que los parámetros obtenidos no tienen una interpretación práctica lo que no nos permite saber los pesos de la red o los valores de activación de los nodos.

La representación matemática de un perceptrón simple con función lineal que es equivalente a un modelo de regresión lineal es:

$$y_{pk} = \theta_k + \sum_{i=1}^n w_{ik} x_{pi}$$

Donde  $\theta_k$  es el umbral de la neurona de salida k, la cual actúa similar a una constante, n indica el número de neuronas de entrada,  $w_{ik}$  es el peso de la variable de entrada i y la de la salida k y  $x_{pi}$  es el valor de la neurona de entrada i y el patrón de entrada k.

En caso de que haya más de una variable salida la red se vuelve en una modelo de regresión multivariado.

Un perceptrón simple con función de activación logística es similar al modelo de regresión logística y matemáticamente su función es la siguiente:

$$y_{pk} = \frac{1}{1 + e^{(-\theta_k + \sum_{i=1}^n w_{ik} x_{pi})}}$$

Para el modelo de perceptrón simple con función umbral equivalente al análisis discriminante lineal la función que lo expresa es:

$$y_{pk} = \frac{e^{l_{pk}}}{\sum_{i=1}^n e^{l_{pk}}}$$



Donde  $l_{pk}$  es la entrada neta que recibe la neurona de salida k.

Por Kolmogorov, sabemos que una función continua de diferentes variables puede ser representada por la concatenación de varias funciones continuas de una misma variable. Lo que indica que un perceptrón que contenga mínimo una capa oculta con suficientes unidades no lineales tiene la capacidad de poder aprender virtualmente cualquier tipo de relación siempre que esta pueda ser aproximada en términos de una función continua.

El modelo perceptrón multicapa auto asociativo equivalente a un análisis de componentes principales está compuesto por una N neuronas de entrada y M neuronas de salida con  $M < N$ . El peso de cada neurona M es el vector propio y la modificación de sus pesos se realiza iterativamente mediante la siguiente fórmula:

$$\Delta w_{ik} = \eta y_{pk} (x_{pi} - y_{pk} w_{ik})$$

Donde la  $\eta$  representa la tasa de aprendizaje y la variable  $y_{pk} = \sum_{i=1}^n w_{ik} x_{pi}$  es la salida de la neurona k para el patrón p y el valor propio  $\lambda$  asociado al vector propio W se calcula con la siguiente fórmula:

$$\lambda = W^T C W$$

Donde C es la matriz de correlaciones.

La ventaja de estos modelos es que mediante un proceso de aprendizaje se pueden sintetizar algoritmos, si no se conocen todos los detalles matemáticos el modelo no se ve afectado y es extremadamente útil para problemas no lineales, sin embargo también tiene desventajas tales como un mayor tiempo de realización con muchas pruebas, el método debe ser entrenado para cada problema particular y el mayor es el que se conoce como efecto Caja Negra donde no se puede apreciar el peso de cada variable en el resultado.

## 2.4 Árboles de Decisión

Es un método analítico que permite obtener una mejora en la toma de decisiones y la eliminación del error humano a través de una representación esquemática de todas las

alternativas disponibles, para lo cual es necesario tener opciones bien definidas, que sean cuantificables y un objetivo claro para que pueda ser llevado a cabo.

Entre los beneficios de este tipo de análisis es que nos permite evaluar consecuencias, implementar variables de riesgo, hacer análisis con variables discretas y/o continuas, hacer predicciones y cálculos de probabilidades.

En la Tabla 2.1 se puede observar a más detalle los cuatro componentes principales de este modelo para poder ser utilizado en decisiones económicas.

En la Figura 2.1 se puede observar así mismo, su representación gráfica.

Componentes	Concepto
Alternativas de Decisión	Son los posibles cursos de acción para el decisor
Eventos Probabilísticos	Son aquellos Eventos que el decisor no controla, pero que deben ser tomados en cuenta para evaluar las alternativas de acción.
Información Económica	Son las consecuencias económicas de las posibles acciones del decisor.
Secuencia del Proceso	<p>Orden en el que se producen y la forma como están relacionadas las decisiones (nodos de decisión) y los eventos probabilísticos (nodos de estado)</p> <ul style="list-style-type: none"> <li>·Nodos de Decisión: Cuadrados/ Rectángulos.</li> <li>·Nodos de Estado (Incertidumbre): Círculos.</li> <li>·Ramas del Árbol: segmento que conecta los nodos.</li> <li>·Probabilidades: se colocan sobre cada una de las ramas que parten de un nodo de estado, de acuerdo con la probabilidad del evento asociado con la rama específica. La suma de las probabilidades de las ramas que parten de un mismo nodo debe ser igual a 1.</li> <li>·Valores económicos: se suelen colocar bajo la rama en la cual se produce un ingreso o un egreso. Al extremo de los nodos terminales debe figurar el resultado de los ingresos y egresos que tuvieron lugar a lo largo del camino.</li> </ul>

Tabla 2.1 Principales componentes de un árbol de decisión.

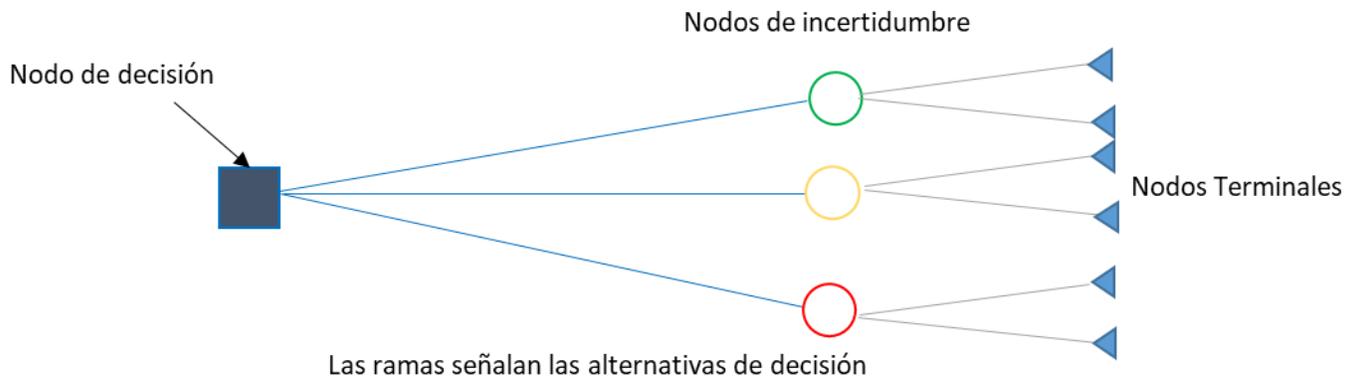


Figura 2.1 Esquema de un árbol de decisión.



## 2.5 Análisis de Componentes Principales (ACP)

Es un método multivariante que nos permite analizar patrones en nuestros datos resaltando las similitudes y diferencias existentes entre ellos, reducir la dimensión de nuestros datos (aunque con alguna pérdida de información) y poder construir un nuevo conjunto de datos no correlacionados entre ellos. (Marín, s.f., Tema 3:1)

Las bases estadísticas que ocupa este modelo son la covarianza, los eigenvalores y los eigenvectores. La covarianza es una medida que trabaja sobre dos dimensiones la cual se ocupa para saber la relación que comparten las variables, pudiendo no estar relacionadas entre ellas (cuando el valor de las covarianzas es cero) y si lo están, ya sea que incrementen conjuntamente sus valores o que mientras una aumente la otra decrece. Su fórmula es la siguiente:

$$cov(X, Y) = \frac{\sum_{i=0}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)}$$

En este modelo comúnmente se ocupan más de dos variables o dos dimensiones por lo que se ocupa la matriz de covarianzas:

$$C^{n*n} = (C_{ij}, C_{ij} = cov(Dim_i, Dim_j))$$

Donde  $C^{n*n}$  es una matriz con n columnas y n filas de dimensión n. Cada entrada de la matriz es la covarianza entre dos dimensiones con la diagonal siendo la varianza de la dimensión donde se encuentra y también esta matriz es simétrica. Los valores propios y vectores propios también conocidos como eigenvalores y eigenvectores<sup>6</sup> se pueden definir de la siguiente manera:

Si A es una matriz nxn, hay vectores x diferentes de cero en  $R^n$  tales que Ax sea un múltiplo escalar de x. El escalar denotado por  $\lambda$  se llama valor propio de la matriz A y el vector x diferente de cero se llama vector propio de A.

---

<sup>6</sup> Proveniente del término alemán eigenwert utilizada por primera vez en matemáticas por el matemático alemán David Hilbert.



Para estudiar las relaciones que se presentan entre  $p$  variables correlacionadas (que miden información común) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tenga repetición o redundancia en la información) llamado conjunto de componentes principales.

Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra. De modo ideal, se buscan  $m < p$  variables que sean combinaciones lineales de las  $p$  originales y que estén incorreladas, recogiendo la mayor parte de la información o variabilidad de los datos.

Si las variables originales están incorreladas de partida, entonces no tiene sentido realizar un análisis de componentes principales.

El análisis de componentes principales es una técnica matemática que no requiere la suposición de normalidad multivariante de los datos, aunque si esto último se cumple se puede dar una interpretación más profunda de dichos componentes.

Así que para calcular los componentes se considera una serie de variables  $(x_1, x_2, \dots, x_p)$  sobre un grupo de objetos o individuos y se trata de calcular, a partir de ellas, un nuevo conjunto de variables  $y_1, y_2, \dots, y_p$ , incorreladas entre sí, cuyas varianzas vayan decreciendo progresivamente.

Cada  $y_j$  (donde  $j = 1, \dots, p$ ) es una combinación lineal de las  $x_1, x_2, \dots, x_p$  originales, es decir:

$$y_j = a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jp}x_p = a'_j x$$

Siendo  $a'_j = (a_{1j}, a_{2j}, \dots, a_{pj})$  un vector de constantes, y



$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$$

Obviamente, si lo que queremos es maximizar la varianza, como veremos luego, una forma simple podría ser aumentar los coeficientes  $a_{ij}$ . Por ello, para mantener la ortogonalidad de la transformación se impone que el módulo del vector  $a'_j = (a_{1j}, a_{2j}, \dots, a_{pj})$  sea

$$a'_j a_j = \sum_{k=1}^p a_{kj}^2 = 1$$

El primer componente se calcula eligiendo  $a_1$  de modo que  $y_1$  tenga la mayor varianza posible, sujeta a la restricción de que  $a'_1 a_1 = 1$ . El segundo componente principal se calcula obteniendo  $a_2$  de modo que la variable obtenida,  $y_2$  esté incorrelada con  $y_1$ . Del mismo modo se eligen  $y_1, y_2, \dots, y_p$  incorrelados entre sí, de manera que las variables aleatorias obtenidas vayan teniendo cada vez menor varianza.

Queremos elegir  $a_1$  de modo que se maximice la varianza de  $y_1$  sujeta a la restricción de que  $a'_1 a_1 = 1$

$$Var(y_1) = Var(a'_1 x) = a'_1 \Sigma a_1$$

El método habitual para maximizar una función de varias variables sujeta a restricciones es el método de los multiplicadores de Lagrange.

El problema consiste en maximizar la función  $a'_1 \Sigma a_1$  sujeta a la restricción  $a'_1 a_1 = 1$ . Se puede observar que la incógnita es precisamente  $a_1$  (el vector desconocido que nos da la combinación lineal óptima).

Así, construyo la función  $L$ :

$$L(a_1) = a'_1 \Sigma a_1 - \lambda(a'_1 a_1 - 1)$$

Y busco el máximo, derivando e igualando a 0:

$$\frac{\partial L}{\partial a_1} = 2\Sigma a_1 - 2\lambda a_1 = 0 \Rightarrow$$



$$(\Sigma - \lambda I)a_1 = 0$$

Esto es, en realidad, un sistema lineal de ecuaciones. Por el teorema de Roché-Frobenius, para que el sistema tenga una solución distinta de 0 la matriz  $(\Sigma - \lambda I)$  tiene que ser singular. Esto implica que el determinante debe ser igual a cero:

$$|\Sigma - \lambda I| = 0$$

y de este modo,  $\lambda$  es un autovalor de  $\Sigma$ . La matriz de covarianzas  $\Sigma$  es de orden  $p$  y si además es definida positiva, tendrá  $p$  auto valores distintos,  $\lambda_1, \lambda_2, \dots, \lambda_p$  tales que, por ejemplo,  $\lambda_1 > \lambda_2 > \dots > \lambda_p$ .

Se tiene que, desarrollando la expresión anterior,

$$(\Sigma - \lambda I)a_1 = 0$$

$$\Sigma a_1 - \lambda I a_1 = 0$$

$$\Sigma a_1 = \lambda I a_1$$

Entonces,

$$\begin{aligned} Var(y_1) &= Var(a'_1 x) = a'_1 \Sigma a_1 = a'_1 \lambda I a_1 \\ &= \lambda a'_1 a_1 = \lambda \cdot 1 = \lambda \end{aligned}$$

Luego, para maximizar la varianza de  $y_1$  se tiene que tomar el mayor autovalor, digamos  $\lambda_1$ , y el correspondiente autovector  $a_1$ . En realidad,  $a_1$  es un vector que nos da la combinación de las variables originales que tiene mayor varianza, esto es, si

$a'_1 = (a_{11}, a_{12}, \dots, a_{1p})$ , entonces

$$y_1 = a'_1 x = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p$$

El segundo componente principal, digamos  $y_2 = a'_2 x$ , se obtiene mediante un argumento parecido. Además, se quiere que  $y_2$  esté incorrelado con el anterior componente  $y_1$ , es decir,  $Cov(y_2, y_1) = 0$ . Por lo tanto:



$$\begin{aligned} \text{Cov}(y_2, y_1) &= \text{Cov}(a'_2 x, a'_1 x) = \\ &= a'_2 \cdot E[(x - \mu)(x - \mu)'] \cdot a_1 = \\ &= a'_2 \Sigma a_1 \end{aligned}$$

Es decir, se requiere que  $a'_2 \Sigma a_1 = 0$

Como se tenía que  $\Sigma a_1 = \lambda a_1$ , lo anterior es equivalente a

$$a'_2 \Sigma a_1 = a'_2 \lambda a_1 = \lambda a'_2 a_1 = 0$$

Esto equivale a que  $a'_2 a_1 = 0$ , es decir, que los vectores sean ortogonales.

De este modo, tendremos que maximizar la varianza de  $y_2$ , es decir  $a_2 \Sigma a_2$ , sujeta a las siguientes restricciones

$$\begin{aligned} a'_2 a_2 &= 1 \\ a'_2 a_1 &= 0 \end{aligned}$$

Se toma la función

$$L(a_2) = a'_2 \Sigma a_2 - \lambda(a'_2 a_2 - 1) - \delta a'_2 a_1$$

Y se deriva

$$\frac{\partial L(a_2)}{\partial a_2} = 2 \Sigma a_2 - 2 \lambda a_2 - \delta a_1 = 0$$

Si se multiplica por  $a'_1$ , entonces

$$2 a'_1 \Sigma a_2 - \delta = 0$$

Porque

$$\begin{aligned} a'_1 a_2 &= a'_2 a_1 = 0 \\ a'_1 a_1 &= 1 \end{aligned}$$

Luego



$$\delta = 2a'_1 \Sigma a_2 = 2a'_2 \Sigma a_1 = 0$$

Ya que  $Cov(y_2, y_1) = 0$ . De este modo,  $\frac{\partial L(a_2)}{\partial a_2}$  queda finalmente como:

$$\frac{\partial L(a_2)}{\partial a_2} = 2\Sigma a_2 - 2\lambda a_2 - \delta a_1 = 2\Sigma a_2 - 2\lambda a_2 = (\Sigma - \lambda I)a_2 = 0$$

Usando los mismos razonamientos que antes, elegimos  $\lambda$  como el segundo mayor autovalor de la matriz  $\Sigma$  con su autovector asociado  $a_2$ .

Los razonamientos anteriores se pueden extender, de modo que al  $j$  –ésimo componente le correspondería el  $j$  –ésimo autovalor.

Entonces todos los componentes  $y$  (en total  $p$ ) se pueden expresar como el producto de una matriz formada por los autovectores, multiplicada por el vector  $x$  que contiene las variables originales  $x_1, x_2, \dots, x_p$

$$y = Ax$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}, \quad A = \begin{pmatrix} a_{11}, a_{12} \dots a_{1p} \\ a_{21}, a_{22} \dots a_{2p} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ a_{p1}, a_{p2} \dots a_{pp} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

Como

$$Var(y_1) = \lambda_1$$

$$Var(y_2) = \lambda_2$$

...

$$Var(y_p) = \lambda_p$$

La matriz de covarianza de  $y$  será

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$



Porque  $y_1, \dots, y_p$  se han construido como variables incorreladas. Se tiene que

$$\Lambda = \text{Var}(Y) = A' \text{Var}(X)A = A' \Sigma A$$

O bien

$$\Sigma = A \Lambda A'$$

Ya que  $A$  es una matriz ortogonal (porque  $a'_i a_i = 1$  para todas sus columnas) por lo que  $AA' = I$

Como se pudo observar cada autovalor correspondía a la varianza del componente  $y_i$  que se definía por medio del autovector  $a_i$ , es decir,  $\text{Var}(y_i) = \lambda_i$ .

Al sumar todos los autovalores, tendremos la varianza total de los componentes, es decir:

$$\sum_{i=1}^p \text{Var}(y_i) = \sum_{i=1}^p \lambda_i = \text{traza}(\Lambda)$$

Ya que la matriz  $\Lambda$  es diagonal.

Pero, por las propiedades del operador traza,

$$\text{traza}(\Lambda) = \text{traza}(A' \Sigma A) = \text{traza}(\Sigma A' A) = \text{traza}(\Sigma)$$

Porque  $AA' = I$  al ser  $A$  ortogonal, con lo cual

$$\text{traza}(\Lambda) = \text{traza}(\Sigma) = \sum_{i=1}^p \text{Var}(x_i)$$

Es decir, la suma de las varianzas de las variables originales y la suma de las varianzas de las componentes son iguales. Esto permite hablar del porcentaje de varianza total que recoge un componente principal:



$$\frac{\lambda_i}{\sum_{i=1}^p \lambda_i} = \frac{\lambda_i}{\sum_{i=1}^p \text{Var}(x_i)}$$

Así, también se podrá expresar el porcentaje de variabilidad recogido por los primeros  $m$  componentes:

$$\frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \text{Var}(x_i)}$$

Donde  $m < p$ .

En la práctica, al tener en principio  $p$  variables, nos quedaremos con un número mucho menor de componentes que recoja un porcentaje amplio de la variabilidad total  $\sum_{i=1}^p \text{Var}(x_i)$ .

Los componentes se calcularán sobre las variables originales estandarizadas, es decir, variables con media 0 y varianza 1. Esto equivale a tomar los componentes principales, no de la matriz de covarianzas sino de la matriz de correlaciones (en las variables estandarizadas coinciden las covarianzas y las correlaciones).

Así, los componentes son autovectores de la matriz de correlaciones y son distintos de los de la matriz de covarianzas. Si se actúa así, se da igual importancia a todas las variables originales.

En la matriz de correlaciones todos los elementos de la diagonal son iguales a 1. Si las variables originales están tipificadas, esto implica que su matriz de covarianzas es igual a la de correlaciones, con lo que la variabilidad total (la traza) es igual al número total de variables que hay en la muestra. La suma total de todos los autovalores es  $p$  y la proporción de varianza recogida por el autovector  $j$ -ésimo (componente) es

$$\frac{\lambda_j}{p}$$



Como se tiene que  $x = A^{-1}y = A'y$ , por ser  $A$  ortogonal, entonces

$$Cov(y_j, x_i) = Cov\left(y_j \sum_{k=1}^p a_{ik}y_k\right) = a_{ij}Var(y_j) = \lambda_i a_{ij}$$

Donde  $y_j$  es el factor  $j$  –ésimo y  $x_i$  es la variable original  $i$  –ésima.

Si suponemos que las variables originales están estandarizadas:

$Var(x_i) = 1$  para  $i = 1, \dots, p$ , entonces

$$Cor(y_j, x_i) = \frac{\lambda_j a_{ij}}{1 \cdot \lambda_j^{1/2}} = \lambda_j^{1/2} a_{ij}$$

De este modo, la matriz de correlaciones entre  $y$  y  $x$  es:

$$Cor(y, x) = \Lambda^{1/2} A' = F'$$

Con lo que la matriz factorial también mide las correlaciones entre las variables originales estandarizadas y los nuevos factores.

Como las variables originales  $x_1, \dots, x_p$  están incorreladas, entonces carece de sentido calcular unos componentes principales. Si se hiciera, se obtendrían las mismas variables pero reordenadas de mayor a menor varianza. Para saber si  $x_1, \dots, x_p$  están correlacionadas, se puede calcular la matriz de correlaciones aplicándose posteriormente el test de esfericidad de Bartlett. El cálculo de los componentes principales de una serie de variables  $x_1, \dots, x_p$  depende normalmente de las unidades de medida empleadas. Si transformamos las unidades de medida, lo más probable es que cambien a su vez los componentes obtenidos. Una solución frecuente es usar variables  $x_1, \dots, x_p$  tipificadas. Con



---

ello, se eliminan las diferentes unidades de medida y se consideran todas las variables implícitamente equivalentes en cuanto a la información recogida.



## Capítulo 3 Construcción del Modelo de Calificación

Como se comentó en el primer capítulo se pueden aplicar distintos modelos para poder administrar el riesgo de crédito. Es por ello, que se ha considerado para este caso aplicar un modelo de puntuación conocido como *Credit Scoring*. El cual es un método estadístico que analiza las características y desempeño de los clientes existentes, para poder predecir el desempeño de estos en un futuro con base en qué tan similares son. Otra ventaja es que facilitará en la institución un análisis rápido del cliente en base a sus razones financieras y adicionada elementos cualitativos para poder clasificarlos en un rango, los demás modelos solo ofrecen una probabilidad de quiebra sin tomar en cuenta las características del colateral del financiamiento.

Este tipo de modelo tiene como objetivo optimizar la toma de decisión con respecto a otorgar y clasificar un crédito, así como, facilitar su correcta administración una vez ya tomado.

### 3.1 Obtención y clasificación de la información de los clientes

Se consideró como referencia para la actualización del modelo de calificación para la medición del riesgo crediticio de la cartera empresarial una base de datos proporcionada por una Institución de Banca de Desarrollo con más de 350 empresas tanto en cartera vencida como en cartera vigente con información obtenida de sus estados financieros trimestrales con un horizonte de tiempo de más de 5 años.

Dicha base contiene el código de cliente, el saldo actual del crédito, el monto original, el plazo, tipo de crédito (revolvente o simple), en caso de ser revolvente la utilización, indicador de en qué cartera se encuentra (vigente o vencida), así como los cambios que ha tenido el crédito en los últimos 12 meses, el sector, entidad, antigüedad así como sus estados financieros trimestrales.



El modelo propuesto se basa en un análisis cuantitativo de diversas razones financieras calculadas a partir de los estados financieros dividiéndose en Capacidad de Pago, Rentabilidad, Tamaño de Ventas, Liquidez, Productividad y Endeudamiento.

### 3.2 Variables para utilizar

Para de **capacidad de pago** se propusieron las siguientes variables:

- Capacidad de pago, medida como EBITDA/Gastos Financieros para medir el desempeño operativo de una compañía en comparación con sus gastos financieros. El EBITDA se deriva de las siglas en inglés de Earnings Before Interests, Taxes, Depreciations and Amortizations el cual indica las ganancias que obtiene la empresa antes de intereses, impuestos depreciaciones y amortizaciones.
- Cobertura del servicio de la deuda, el cual se calcula como EBITDA/Gastos Financieros y Deuda a Corto Plazo. Este indicador muestra la capacidad de pago de la empresa asumiendo la deuda que posea a un periodo menor a un año.
- Capacidad de pago con depreciación y amortización, medida como EBIT/Gastos Financieros para medir el desempeño operativo de una compañía en comparación con sus gastos financieros. Este indicador a diferencia del EBITDA no toma en cuenta las depreciaciones y amortizaciones, que algunos especialistas califican como gastos ficticios.
- Cobertura del servicio de la deuda con depreciación y amortización, calculada como EBIT/Gastos Financieros y Deuda a Corto Plazo. Este indicador muestra la capacidad de pago de la empresa asumiendo la deuda que posea a un periodo menor a un año.
- Pasivo a corto plazo/ EBITDA + Gastos financieros, esta razón se calcula con el pasivo a corto plazo entre el EBITDA aunado a los gastos financieros e indica que tanto representa de la ganancia incluyendo los gastos financieros las obligaciones exigibles en el corto plazo.



En **rentabilidad** se propusieron las siguientes razones:

- Retorno sobre Activos, conocido normalmente como ROA (*“Return on Assets”*) y se calcula como el cociente de las utilidades netas y el activo total de la empresa, y mide la eficiencia de la empresa a través del activo.
- Retorno sobre el Patrimonio, conocido normalmente como ROE (*“Return on Equity”*) y se calcula como el cociente de las utilidades netas y el capital de la empresa, mide el rendimiento del capital invertido con respecto a las ganancias.
- Retorno sobre la Inversión, conocido normalmente como ROI (*“Return on Investment”*) y se calcula como el cociente de las utilidades netas menos gastos de administración y costos de venta y operación sobre los gastos de administración y costos de venta y operación, mide el beneficio obtenido a través de lo gastado.

Para las razones de **endeudamiento** se propusieron las siguientes:

- Endeudamiento, este indicador se mide como Pasivo Total entre Activo Total e indica el tamaño proporcional de la deuda respecto al activo, se puede interpretar como que tanto se financia la empresa con recursos internos o externos.
- Apalancamiento, medido como Pasivo a Corto Plazo entre Capital Contable e indica el grado de dependencia de la empresa respecto a sus acreedores.
- Endeudamiento respecto a sus ganancias, medido como pasivo total entre su EBITDA, este indicador muestra en que tantos años podría cubrir toda su deuda utilizando las ganancias constantes.
- Cobertura de intereses, medido como EBIT entre gastos de interés, esta razón se determina con base en las utilidades antes de impuestos ya que los intereses son deducibles de estos. Muestra hasta donde las utilidades pueden soportar el cubrir los intereses.



Para **liquidez** se propusieron las siguientes razones:

- Liquidez o índice de circulante se calcula como activos circulantes entre pasivos circulantes, esta razón da el nivel de seguridad que tienen los acreedores en el corto plazo.
- Prueba Ácida se calcula como activos circulantes menos inventarios entre pasivos circulantes, esta razón es parecida a la anterior solo que toma los activos más líquidos y en caso de que la empresa a evaluar sea prestadora de servicios será el mismo valor que la razón anterior.
- Capital de trabajo esta razón se calcula como los activos circulantes menos pasivos circulantes y mide la disponibilidad con que cuenta la empresa en el corto plazo para hacer frente a sus obligaciones.

Para **ventas** se propusieron los siguientes indicadores:

- Utilidad de operación entre ventas netas, se calcula con los ingresos operacionales, restándosele los costos y gastos operacionales sobre las ventas netas, esta razón mide la utilidad de la actividad principal de la empresa.
- Crecimiento de ventas, se calcula dividiendo las ventas del periodo actual entre las ventas del periodo anterior y se le resta una unidad, este indicador mide que tanto mejoraron o disminuyeron las ventas sobre el periodo anterior.
- Rotación de inventario, este indicador se calcula dividiendo el costo de las mercancías vendidas en un determinado periodo sobre el promedio de los inventarios sobre dicho periodo, indicándonos el volumen que se tiene de ventas.
- Rotación de ventas, se calcula mediante las ventas dividido entre el activo total e indica el número de veces que a un cierto nivel de ventas se utilizan nuestros activos.
- Impacto de los gastos de administración y ventas, el cual se calcula mediante la suma de los gastos de administración y ventas divididos entre las ventas netas indicando que tanto porcentaje de las ventas se usa en este tipo de gastos.



### 3.3 Reducción a las variables más discriminantes

Una vez definidas las variables que pueden ayudar a identificar el riesgo de que una empresa presente un incumplimiento se verá qué metodología se puede utilizar.

El análisis discriminante requiere para su utilización ciertos supuestos ya enumerados anteriormente y de los cuales no se cumplen principalmente dos: no contamos con una variable categórica y las variables no siguen una distribución normal multivariante.

Para poder aplicar alguno de los modelos de regresión de los que se habló anteriormente se necesita una variable dicotómica al igual que en el análisis anterior, generalmente para este tipo de modelos se utiliza el incumplimiento, indicando si el acreditado se encuentra en este estado o no. La base que vamos a utilizar cuenta con muy pocos acreditados en estado de incumplimiento por lo tanto se complica la aplicación de alguno de los tres modelos de regresión explicados.

Para la aplicación de un modelo de redes neuronales se necesita ocupar un software especializado, así como una calibración constante, por lo que se pudiese aplicar una vez ya implementado para llevarlo a la par y observar los resultados.

El método de árboles de decisión necesita que las variables tengan una cierta jerarquía para poder ver algún posible curso de acción para llegar a una decisión, por lo que se podrá ocupar este método una vez se haya llevado a cabo el análisis cuantitativo en el análisis cualitativo.

Por lo tanto, el modelo a implementar para poder discriminar a las variables a ocupar fue el modelo de análisis de componentes principales, las variables se analizaron mediante este método utilizando un software estadístico llamado XLSTAT.



En el primer análisis de los datos se puede observar en la matriz de correlaciones que varias variables tienen una correlación positiva muy cercana a la unidad lo que nos indica que se puede eliminar cualquier variable de las que presentan este tipo de correlación sin afectar el resultado.

La matriz de correlaciones del modelo se puede observar en la Tabla 2.2, en esta se encuentran marcadas algunas razones que presentan una muy alta correlación entre ellas.

El primer par de razones que presentan una alta correlación, en este caso, 0.999 son liquidez y capital de trabajo, de ambas se seleccionó la primera debido a su fácil interpretación y mayor aceptación dentro del ámbito financiero.

El siguiente par de razones que muestran una correlación negativa muy alta entre ellas de -0.996 son Utilidad de la Operación entre Ventas e Impacto de los Gastos de Administración y Ventas lo que indica que son inversamente proporcionales, es decir, que mientras una incrementa la otra disminuye, de estas dos se eligió la Utilidad entre Ventas ya que nos muestra el ingreso de nuestra actividad principal.

Las siguientes razones que muestran una correlación positiva son EBITDA/GF con EBIT/GF (0.972) y al mismo tiempo las dos con Cobertura de intereses (EBITDA/GF 0.965 y con EBIT/GF 0.994), en este caso se seleccionó la variable EBITDA/ GF ya que este indicador se utiliza en varios créditos como un covenant financiero (son razones que se utilizan a la hora de otorgar un crédito, las cuales son usadas como límites para obtener ciertas tasas).

Otra pareja de razones que presentan alta correlación son EBITDA/ (GF +DCP) con EBIT/ (GF +DCP) de 0.978, al igual que el caso anterior se usará EBITDA/ (GF +DCP) por la misma razón. Por último, las razones que vuelven a presentar alta correlación positiva son Endeudamiento respecto a sus ganancias y Pasivo Total entre EBITDA con 1, se seleccionó la última por la utilización del EBITDA.



Variables	Liquidez	Prueba Ácida	Capital de Trabajo	Crecimiento en ventas	Utilidad de la operación	Rotación de inventarios	Rotación de ventas	Impacto de los GA y V	ROA	ROE	ROI	Apalancamiento	Endeudamiento	Endeudamiento Ganancias	Cobertura de intereses	EBITDA/GF	EBITDA/GF+DCP	EBIT/GF	EBIT/GF+DCP	PT/EBITDA	PCC/EBITDA-GF
Liquidez	1	0.731	0.999	0.001	-0.133	-0.126	-0.144	0.121	0.003	0.044	0.049	-0.096	-0.185	0.112	0.038	0.078	-0.002	0.053	-0.006	0.112	0.153
Prueba Ácida	0.731	1	0.811	-0.009	-0.046	-0.076	-0.196	0.064	-0.024	0.073	0.040	-0.102	-0.182	0.111	0.055	0.099	0.018	0.066	0.012	0.111	0.139
Capital de Trabajo	0.999	0.811	1	0.001	-0.130	-0.126	-0.147	0.122	0.004	0.043	0.051	-0.094	-0.185	0.111	0.037	0.078	0.001	0.054	-0.003	0.111	0.151
Crecimiento en ventas	0.001	-0.009	0.001	1	0.073	-0.027	0.051	-0.034	0.076	0.133	0.008	-0.001	0.027	0.103	0.015	0.020	0.017	0.017	0.027	0.103	0.065
Utilidad de la operación	-0.133	-0.046	-0.130	0.073	1	0.014	0.038	-0.995	0.226	0.256	0.046	-0.015	-0.056	0.102	0.213	0.207	0.278	0.215	0.297	0.102	0.004
Rotación de inventarios	-0.126	-0.076	-0.126	-0.027	0.014	1	0.131	0.037	-0.041	-0.021	-0.011	0.004	0.069	-0.104	-0.318	-0.325	-0.245	-0.323	-0.229	-0.104	-0.004
Rotación de ventas	-0.144	-0.196	-0.147	0.051	0.038	0.131	1	0.049	0.277	0.087	0.001	0.169	0.060	-0.003	-0.215	-0.238	-0.226	-0.215	-0.233	-0.003	-0.009
Impacto de los GA y V	0.121	0.064	0.122	-0.034	-0.995	0.037	0.049	1	-0.190	-0.204	-0.024	0.037	0.012	-0.165	-0.269	-0.264	-0.328	-0.273	-0.351	-0.165	-0.006
ROA	0.003	-0.024	0.004	0.076	0.226	-0.041	0.277	-0.190	1	0.291	0.212	-0.037	-0.433	0.166	0.264	0.250	0.292	0.269	0.309	0.166	0.087
ROE	0.044	0.073	0.043	0.133	0.256	-0.021	0.087	-0.204	0.291	1	-0.082	-0.559	-0.070	0.057	0.112	0.113	0.100	0.107	0.106	0.057	-0.029
ROI	0.049	0.040	0.051	0.008	0.046	-0.011	0.001	-0.024	0.212	-0.082	1	0.007	-0.191	0.128	0.133	0.161	0.167	0.156	0.173	0.128	0.156
Apalancamiento	-0.096	-0.102	-0.094	-0.001	-0.015	0.004	0.169	0.037	-0.037	-0.559	0.007	1	0.166	0.081	0.013	0.013	0.001	0.012	-0.008	0.081	0.114
Endeudamiento	-0.185	-0.182	-0.185	0.027	-0.056	0.069	0.060	0.012	-0.433	-0.070	-0.191	0.166	1	-0.087	-0.038	-0.009	-0.026	-0.040	-0.050	-0.087	-0.086
Endeudamiento G	0.112	0.111	0.111	0.103	0.102	-0.104	-0.003	-0.165	0.166	0.057	0.128	0.081	-0.087	1	0.335	0.390	0.206	0.348	0.193	1.000	0.472
Cobertura de intereses	0.038	0.055	0.037	0.015	0.213	-0.318	-0.215	-0.269	0.264	0.112	0.133	0.013	-0.038	0.335	1	0.965	0.740	0.994	0.760	0.335	0.170
EBITDA/GF	0.078	0.099	0.078	0.020	0.207	-0.325	-0.238	-0.264	0.250	0.113	0.161	0.013	-0.009	0.390	0.965	1	0.747	0.972	0.735	0.390	0.196
EBITDA/GF+DCP	-0.002	0.018	0.001	0.017	0.278	-0.245	-0.226	-0.328	0.292	0.100	0.167	0.001	-0.026	0.206	0.740	0.747	1	0.743	0.978	0.206	0.029
EBIT/GF	0.053	0.066	0.054	0.017	0.215	-0.323	-0.215	-0.273	0.269	0.107	0.156	0.012	-0.040	0.348	0.994	0.972	0.743	1	0.758	0.348	0.170
EBIT/GF+DCP	-0.006	0.012	-0.003	0.027	0.297	-0.229	-0.233	-0.351	0.309	0.106	0.173	-0.008	-0.050	0.193	0.760	0.735	0.978	0.758	1	0.193	0.026
PT/EBITDA	0.112	0.111	0.111	0.103	0.102	-0.104	-0.003	-0.165	0.166	0.057	0.128	0.081	-0.087	1.000	0.335	0.390	0.206	0.348	0.193	1	0.472
PCC/EBITDA-GF	0.153	0.139	0.151	0.065	0.004	-0.004	-0.009	-0.006	0.087	-0.029	0.156	0.114	-0.086	0.472	0.170	0.196	0.029	0.170	0.026	0.472	1

Tabla 2.2 Matriz de Correlaciones



Una vez que se eliminaron dichas variables se volvió a analizar la matriz de correlaciones con las 15 restantes, obteniendo los siguientes resultados:

Variables	Liquidez	Prueba Ácida	Crecimiento en ventas	Utilidad de la operación	Rotación de inventarios	Rotación de ventas	ROA	ROE	ROI	Apalancamiento	Endeudamiento	EBITDA/GF	EBITDA/GF+DC	PT/EBITDA	PCC/EBITDA-GF
Liquidez	1	0.731	0.001	-0.133	-0.126	-0.144	0.003	0.044	0.049	-0.096	-0.185	0.078	-0.002	0.112	0.153
Prueba Ácida	0.731	1	-0.009	-0.046	-0.076	-0.196	-0.024	0.073	0.040	-0.102	-0.182	0.099	0.018	0.111	0.139
Crecimiento en ventas	0.001	-0.009	1	0.073	-0.027	0.051	0.076	0.133	0.008	-0.001	0.027	0.020	0.017	0.103	0.065
Utilidad de la operación	-0.133	-0.046	0.073	1	0.014	0.038	0.226	0.256	0.046	-0.015	-0.056	0.207	0.278	0.102	0.004
Rotación de inventarios	-0.126	-0.076	-0.027	0.014	1	0.131	-0.041	-0.021	-0.011	0.004	0.069	-0.325	-0.245	-0.104	-0.004
Rotación de ventas	-0.144	-0.196	0.051	0.038	0.131	1	0.277	0.087	0.001	0.169	0.060	-0.238	-0.226	-0.003	-0.009
ROA	0.003	-0.024	0.076	0.226	-0.041	0.277	1	0.291	0.212	-0.037	-0.433	0.250	0.292	0.166	0.087
ROE	0.044	0.073	0.133	0.256	-0.021	0.087	0.291	1	-0.082	-0.559	-0.070	0.113	0.100	0.057	-0.029
ROI	0.049	0.040	0.008	0.046	-0.011	0.001	0.212	-0.082	1	0.007	-0.191	0.161	0.167	0.128	0.156
Apalancamiento	-0.096	-0.102	-0.001	-0.015	0.004	0.169	-0.037	-0.559	0.007	1	0.166	0.013	0.001	0.081	0.114
Endeudamiento	-0.185	-0.182	0.027	-0.056	0.069	0.060	-0.433	-0.070	-0.191	0.166	1	-0.009	-0.026	-0.087	-0.086
EBITDA/GF	0.078	0.099	0.020	0.207	-0.325	-0.238	0.250	0.113	0.161	0.013	-0.009	1	0.747	0.390	0.196
EBITDA/GF+DCP	-0.002	0.018	0.017	0.278	-0.245	-0.226	0.292	0.100	0.167	0.001	-0.026	0.747	1	0.206	0.029
PT/EBITDA	0.112	0.111	0.103	0.102	-0.104	-0.003	0.166	0.057	0.128	0.081	-0.087	0.390	0.206	1	0.472
PCC/EBITDA-GF	0.153	0.139	0.065	0.004	-0.004	-0.009	0.087	-0.029	0.156	0.114	-0.086	0.196	0.029	0.472	1

Tabla 2.3 Segunda Matriz de Correlaciones

En la Tabla 2.3 se puede observar que las variables ya no presentan una alta correlación entre sí por lo que se procedió a aplicar el Análisis de Componentes Principales, se aplicó el índice de Kaiser-Meyer-Olkin o medida de adecuación muestral para evaluar que el uso de este análisis sea significativo. Se obtuvo un 0.712 y al ser mayor a 0.5 se puede aplicar.

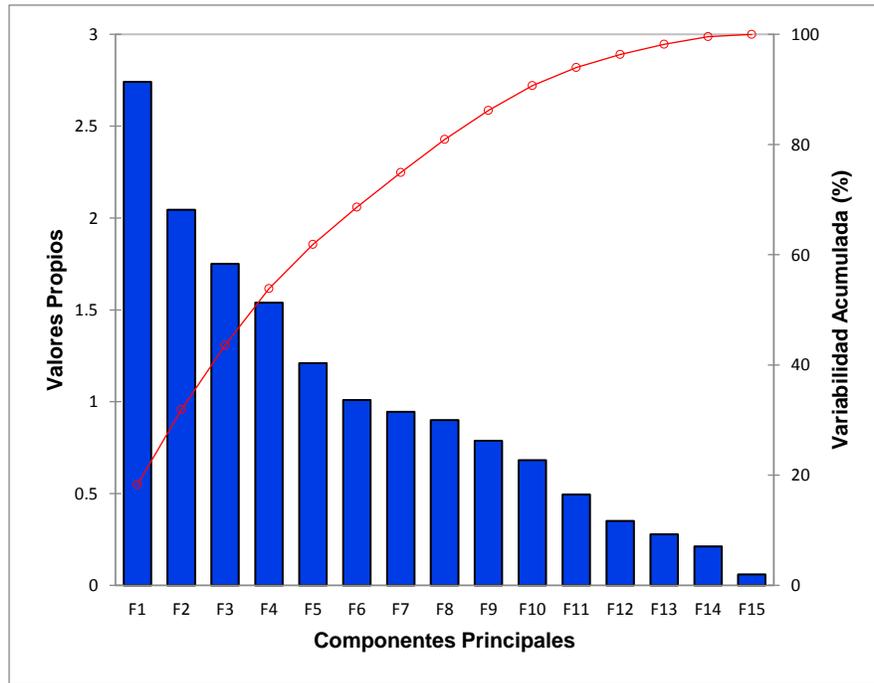


Figura 2.2 Gráfico de Sedimentación

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
Valores Propios	2.741	2.044	1.750	1.538	1.210	1.010	0.944	0.900	0.787	0.682	0.494	0.350	0.278	0.213	0.060
Variabilidad (%)	18.271	31.630	11.665	10.256	8.066	6.730	6.294	5.997	5.248	4.546	3.291	2.336	1.852	1.418	0.399
Acumulada %	18.271	31.901	43.566	53.822	61.888	68.619	74.913	80.909	86.157	90.704	93.995	96.331	98.183	99.601	100.000

Tabla 2.4 Componentes y Variabilidad

Se puede observar que se obtuvieron 15 Componentes, uno por cada variable utilizada en el análisis, para poder calcular el número de factores a extraer existen 4 criterios:

- Criterio de la Raíz Latente. Cualquier componente debe justificar la varianza de por lo menos una variable, es decir, los que posean un valor propio superior a la unidad.
- Criterio a Priori. El analista ya sabe cuántos factores extraer con base en su experiencia

- Criterio de Porcentaje de la Varianza. Aproximación para extraer un porcentaje acumulado específico de la varianza total extraída.
- Criterio de Contraste de Caída. Se basa en ocupar los factores antes de que su cantidad de varianza única domine la varianza común.

Se utilizó el criterio de la raíz latente observándose en la Tabla 2.3 que 6 componentes cumplen con este criterio y aportan una variabilidad acumulada de casi 70% y un 32% en los dos primeros componentes principales.

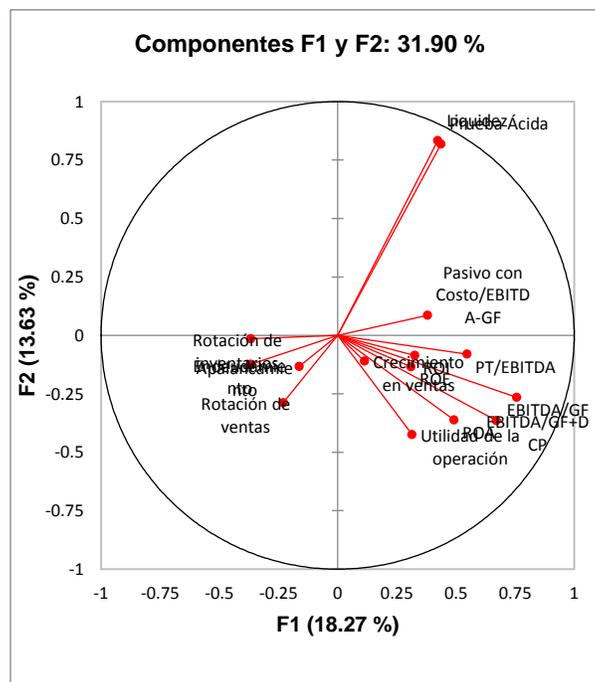


Figura 2.3 Representación en dos ejes de los dos primeros componentes.

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
Liquidez	0.179	<b>0.694</b>	0.003	0.006	0.001	0.063	0.013	0.001	0.009	0.000	0.000	0.000	0.001	0.000	0.030
Prueba Ácida	0.191	<b>0.670</b>	0.005	0.001	0.002	0.052	0.042	0.005	0.002	0.000	0.000	0.000	0.000	0.000	0.029
Crecimiento en ventas	0.013	0.012	0.011	<b>0.372</b>	0.256	0.061	0.228	0.019	0.012	0.013	0.000	0.003	0.000	0.000	0.000
Utilidad de la operación	0.099	0.179	0.045	0.002	0.013	<b>0.252</b>	0.032	0.044	0.103	0.223	0.000	0.002	0.005	0.001	0.000
Rotación de inventarios	0.135	0.000	0.032	0.105	0.010	0.087	0.201	0.098	0.000	<b>0.325</b>	0.002	0.001	0.002	0.001	0.000
Rotación de ventas	0.053	0.083	0.044	0.072	0.010	0.177	0.001	<b>0.342</b>	0.123	0.003	0.002	0.089	0.003	0.000	0.000
ROA	<b>0.242</b>	0.131	0.126	0.163	0.075	0.044	0.002	0.011	0.000	0.059	0.011	0.065	0.069	0.001	0.000
ROE	0.096	0.018	<b>0.558</b>	0.021	0.114	0.004	0.000	0.020	0.026	0.001	0.004	0.052	0.084	0.002	0.000
ROI	0.106	0.007	0.011	0.099	0.169	0.079	0.003	0.199	<b>0.235</b>	0.082	0.003	0.004	0.002	0.000	0.000
Apalancamiento	0.027	0.018	<b>0.498</b>	0.108	0.006	0.180	0.017	0.002	0.022	0.001	0.002	0.061	0.059	0.001	0.000
Endeudamiento	0.136	0.015	0.112	0.116	<b>0.262</b>	0.016	0.043	0.000	0.224	0.002	0.004	0.029	0.041	0.000	0.000
EBITDA/GF	<b>0.573</b>	0.070	0.086	0.090	0.001	0.001	0.002	0.003	0.013	0.032	0.000	0.009	0.008	0.111	0.000
EBITDA/GF+DCP	<b>0.449</b>	0.132	0.039	0.140	0.017	0.012	0.016	0.006	0.009	0.058	0.005	0.027	0.001	0.090	0.000
PT/EBITDA	<b>0.298</b>	0.006	0.082	0.102	0.148	0.055	0.001	0.045	0.004	0.000	0.249	0.003	0.002	0.004	0.000
PCC/EBITDA-GF	0.144	0.007	0.097	<b>0.224</b>	0.126	0.147	0.000	0.022	0.006	0.009	0.213	0.004	0.000	0.001	0.000

Tabla 2.5 Cosenos cuadrados de las variables



Con base en los cosenos cuadrados se aprecia en que ejes se pueden ver mejor las variables que mejor representan la función y le otorgan mayor variabilidad por lo que ahora procederemos a eliminar las razones que no se encuentran en los componentes elegidos.

Por último, vamos a realizar el análisis sobre las razones elegidas, observándose que aún tenemos suficientes razones de cada tipo elegido dando un total de 12 razones.

Variables	Liquidez	Prueba Ácida	Crecimiento en ventas	Utilidad de la operación	ROA	ROE	Apalancamiento	Endeudamiento	EBITDA/GF	EBITDA/GF+DCP	PT/EBITDA	PCC/EBITDA-GF
Liquidez	1	0.731	0.001	-0.133	0.003	0.044	-0.096	-0.185	0.078	-0.002	0.112	0.153
Prueba Ácida	0.731	1	-0.009	-0.046	-0.024	0.073	-0.102	-0.182	0.099	0.018	0.111	0.139
Crecimiento en ventas	0.001	-0.009	1	0.073	0.076	0.133	-0.001	0.027	0.020	0.017	0.103	0.065
Utilidad de la operación	-0.133	-0.046	0.073	1	0.226	0.256	-0.015	-0.056	0.207	0.278	0.102	0.004
ROA	0.003	-0.024	0.076	0.226	1	0.291	-0.037	-0.433	0.250	0.292	0.166	0.087
ROE	0.044	0.073	0.133	0.256	0.291	1	-0.559	-0.070	0.113	0.100	0.057	-0.029
Apalancamiento	-0.096	-0.102	-0.001	-0.015	-0.037	-0.559	1	0.166	0.013	0.001	0.081	0.114
Endeudamiento	-0.185	-0.182	0.027	-0.056	-0.433	-0.070	0.166	1	-0.009	-0.026	-0.087	-0.086
EBITDA/GF	0.078	0.099	0.020	0.207	0.250	0.113	0.013	-0.009	1	0.747	0.390	0.196
EBITDA/GF+DCP	-0.002	0.018	0.017	0.278	0.292	0.100	0.001	-0.026	0.747	1	0.206	0.029
PT/EBITDA	0.112	0.111	0.103	0.102	0.166	0.057	0.081	-0.087	0.390	0.206	1	0.472
PCC/EBITDA-GF	0.153	0.139	0.065	0.004	0.087	-0.029	0.114	-0.086	0.196	0.029	0.472	1

Tabla 2.6 Matriz de Correlación Final

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Valores Propios	2.551	2.006	1.711	1.218	1.143	0.946	0.810	0.552	0.495	0.290	0.216	0.063
Variabilidad (%)	21.258	16.715	14.259	10.149	9.521	7.882	6.752	4.600	4.123	2.420	1.799	0.523
Acumulada %	21.258	37.973	52.232	62.381	71.902	79.784	86.536	91.136	95.259	97.679	99.477	100.000

Tabla 2.7 Componentes y Variabilidad

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Liquidez	0.257	0.604	0.034	0.166	-0.025	0.151	0.043	0.106	0.015	-0.076	-0.024	-0.705
Prueba Ácida	0.270	0.587	0.020	0.201	-0.015	0.164	0.136	0.046	0.022	-0.007	0.001	0.702
Crecimiento en ventas	0.088	-0.057	-0.040	-0.229	0.528	0.702	-0.356	-0.177	-0.061	-0.026	0.025	0.011
Utilidad de la operación	0.232	-0.287	-0.138	0.051	0.020	0.287	0.819	-0.238	-0.031	-0.136	0.080	-0.070
ROA	0.349	-0.177	-0.186	-0.345	-0.358	0.156	-0.119	0.584	0.014	-0.416	0.092	0.055
ROE	0.252	-0.062	-0.539	0.016	0.332	-0.151	0.076	0.369	0.040	0.591	-0.121	-0.026
Apalancamiento	-0.120	-0.090	0.573	-0.098	-0.223	0.383	0.175	0.351	0.085	0.526	-0.073	-0.015
Endeudamiento	-0.234	-0.133	0.208	0.461	0.508	-0.027	0.113	0.517	0.004	-0.363	0.017	0.020
EBITDA/GF	0.449	-0.220	0.214	0.342	-0.021	-0.089	-0.222	-0.075	-0.056	0.181	0.698	-0.021
EBITDA/GF+DCP	0.399	-0.293	0.117	0.417	-0.154	0.055	-0.215	-0.112	-0.170	-0.034	-0.672	0.008
PT/EBITDA	0.353	-0.035	0.318	-0.269	0.272	-0.262	0.040	-0.098	0.719	-0.099	-0.144	0.004
PCC/EBITDA-GF	0.244	0.089	0.348	-0.414	0.274	-0.318	0.143	0.043	-0.661	-0.015	-0.064	0.011

Tabla 2.8 Vectores Propios

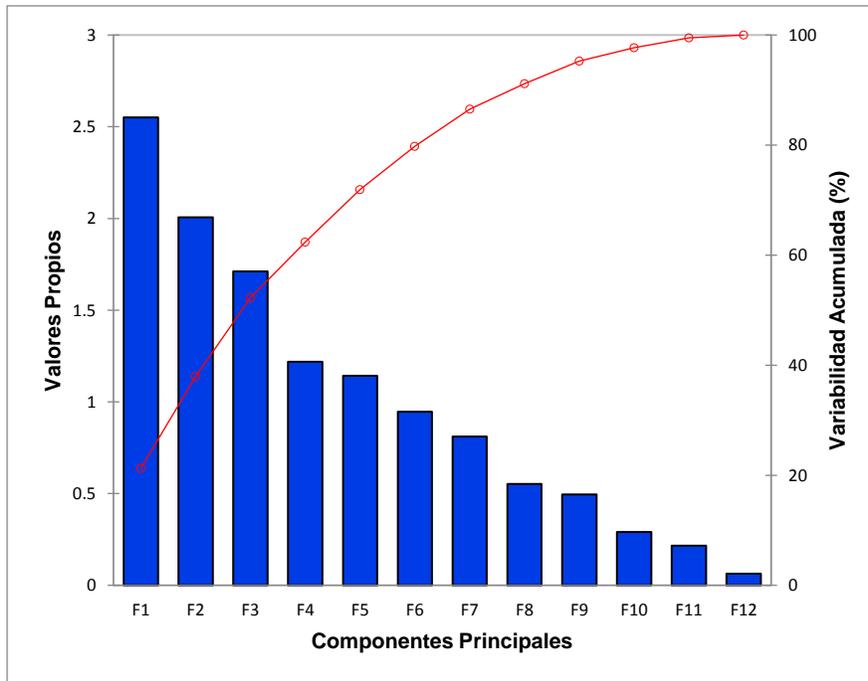


Figura 2.4 Gráfico de Sedimentación Final

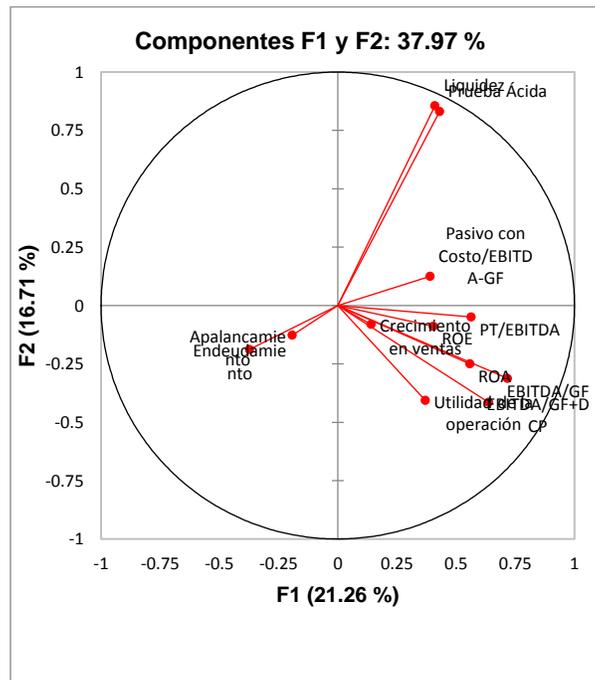


Figura 2.5 Representación en dos ejes de los dos primeros componentes.



Se puede observar que ahora los 5 primeros componentes aportan una variabilidad de casi el 72% y los primeros dos casi el 38%, y que se encuentran razones representativas de cada tipo elegido.

### 3.4 Selección de variables

Después de analizar todas las variables utilizando el análisis de componentes principales se obtuvieron las variables que mejor se ajustaban a la toma del riesgo, las cuales son las siguientes:

- Liquidez
- Prueba Ácida.
- Crecimiento en Ventas.
- Utilidad de la Operación/ Ventas.
- ROA.
- ROE.
- Nivel de Apalancamiento.
- Nivel de Endeudamiento.
- EBITDA/GF.
- EBITDA/GF+DCP.
- PT/EBITDA.
- Pasivo con Costo/EBITDA-GF.

### 3.5 Obtención de la escala de calificación del modelo

Se estableció que para diferenciar la exposición al riesgo habría 9 niveles de calificación distintos para medir el riesgo de crédito. Para la asignación de los intervalos de valores a los 9 intervalos de calificación de riesgo se distribuyeron con base en las medidas estadísticas obtenidas, quedando por lo general en el nivel de riesgo 3 los valores promedio y a partir del nivel 6 se asignaron los valores negativos o muy deteriorados. Para los rangos 1, 2, 4 y



5 se fueron fijando con base en desviaciones estándar. En casos donde los resultados obtenidos de esta forma no se asemejaban a la realidad o diferían mucho por observaciones atípicas que no son tan comunes se ocuparon los percentiles para la asignación de los intervalos.

Adicionalmente a esta calificación obtenida mediante el análisis de indicadores cuantitativos, se tomaron en cuenta aspectos cualitativos del sector, mercado y administración de la empresa que solicita el crédito, así como su experiencia de pago ya sea con la institución o terceros, esta información puede impactar positiva o negativamente la calificación obtenida mediante el análisis financiero en un nivel de riesgo.

Finalmente se puede mejorar la calificación obtenida mediante estos dos análisis únicamente en un nivel de riesgo con base en las garantías que se otorguen y al aforo de las mismas respecto al crédito. El tipo de garantía que se ofrezca establece un nivel mínimo de aforo para que pueda mejorarse la calificación.

Los 9 intervalos posibles de calificación son los siguientes:

1. RIESGO EXCELENTE (R1) - Los indicadores del cliente son excelentes, y cuenta con una posición de liderazgo en el mercado nacional y/o internacional.
2. RIESGO MUY BUENO (R2) - Los indicadores del cliente son muy buenos y participa con una porción de mercado considerable.
3. RIESGO BUENO (R3) - Los indicadores del cliente son buenos y sus indicadores financieros son similares al promedio de la industria.
4. RIESGO REGULAR (R4) - Los indicadores del cliente son regulares con indicadores positivos en general.
5. RIESGO MÍNIMO ACEPTABLE (R5) - Los indicadores del cliente son mínimos, con indicadores financieros por debajo del promedio, resultados mixtos y cercanos al punto de equilibrio.



6. RIESGO DÉBIL (R6) - Los indicadores del cliente son débiles pudiendo presentar pérdidas netas que pudieran ser temporales o bien, está en vías de un mayor deterioro, así como indicadores negativos.
7. RIESGO MUY DÉBIL (R7) - Los indicadores del cliente son muy débiles, con pérdidas importantes y con una alta posibilidad de incumplir en sus pagos en el corto plazo.
8. RIESGO ALTO (R8) - Los indicadores del cliente son insatisfactorios encontrándose en duda la viabilidad de la empresa a causa del deterioro financiero.
9. RIESGO MUY ALTO (R9)- Los indicadores del cliente son nulos y se trata de un negocio que no tiene viabilidad financiera ni de mercado. Está cercana o en quiebra técnica.

Una vez establecidos los intervalos de calificación se consideran como sujetos a crédito si se encuentran dentro de los primeros 5 intervalos de calificación de riesgo y en caso contrario se considera que su nivel de riesgo es muy elevado para aceptarlo y se rechaza automáticamente.

A cada uno de estos 9 niveles se le asignó una probabilidad de incumplimiento para poder calcular su posible pérdida esperada asociada en conjunto con una severidad que se ajusta dependiendo del tipo y aforo de garantía que se tenga para él crédito y exposición.

Para un mejor seguimiento y administración de la cartera crediticia de la institución financiera, esta está subdividida en sectores (Corporativo o Gran Empresa, Comercio, Turismo, Manufacturas y Servicios) para una mejor identificación de las características financieras específicas de cada uno de acuerdo con su nicho de negocio.

Para el cálculo de la probabilidad de incumplimiento existen tres principales formas de calcularlo:

- Mediante un enfoque basado en la experiencia de incumplimiento histórica.
- Mediante un enfoque basado en modelos estadísticos.



- Mediante un enfoque basado en aproximación de una agencia calificadora.

Al poseer información histórica de la cartera a evaluar se utilizó el enfoque basado en la experiencia histórica de incumplimiento. Primero se calcula la frecuencia de incumplimiento de cada nivel de riesgo en un horizonte de tiempo mediante la siguiente fórmula:

$$FI_t = \frac{I_t}{N_t}$$

Donde  $I_t$  es el número de empresas que se encuentra en incumplimiento en el año  $t$  y  $N_t$  el total de empresas que forman parte del nivel de riesgo a evaluar en el año  $t$ .

Con esta fórmula se calcula la frecuencia de incumplimiento total:

$$FIT = \frac{1}{T} \sum_{t=1}^T FI_t$$

Para complementar el cálculo de la probabilidad de incumplimiento se adicionó al mismo también, un enfoque basado en una aproximación de una agencia calificadora, utilizando la probabilidad de incumplimiento con información a 10 años de la Agencia Moody's.

La Probabilidad de Incumplimiento o PI para cada nivel de riesgo en el modelo quedo de la siguiente manera a un horizonte de tiempo de un año:

Indicador	PI
R1	0.1%
R2	0.5%
R3	0.8%
R4	2.4%
R5	4.2%
R6	12.8%
R7	39.1%
R8	65.0%
R9	100%

Tabla 2.9 Probabilidades de Incumplimiento a un Año.



## Capítulo 4 Validación del Modelo

Para una correcta validación del modelo se debe observar el tipo de enfoque que se ocupó para calcular los niveles de riesgo del modelo, esencialmente, hay dos enfoques distintos:

- Los calculados a un determinado punto en el tiempo.
- Los calculados de acuerdo con el ciclo económico.

Los primeros se basan en un ajuste rápido al ambiente económico mientras que el segundo tiende a permanecer constante mediante el ambiente macroeconómico cambia. La principal diferencia de estos dos tipos de enfoque son las variables a las que se les asigna mayor peso, por ejemplo, un enfoque determinado en punto en el tiempo le da mayor peso a la información específica de un acreditado, pudiendo ser esta la línea de negocio, razones financieras o aspectos cualitativos mientras que el otro enfoque le otorga mayor peso a la información macroeconómica pudiendo ser el PIB del sector, la tasa de crecimiento de este, o exportaciones del sector.

El modelo utilizado está basado en el enfoque a un determinado punto en el tiempo al igual que el método regulatorio y las calificaciones de agencias especializadas por lo que se validarán los resultados obtenidos en comparación a estos.

### 4.1 Comparación con modelos de agencias especializadas

Este tipo de agencias por lo general generan dos tipos de calificaciones, de largo y corto plazo, en este trabajo nos centraremos en las de largo plazo debido a que es el tipo de crédito que otorga la institución financiera.

Para estas calificaciones se utilizan también factores cualitativos y cuantitativos comprendiendo los riesgos del negocio, así como los financieros utilizando un perfil



histórico y uno proyectado. Esta calificación posteriormente se sensibiliza con base al perfil de riesgo del sector, la ubicación, el gobierno o estrategia de administración y la estructura del grupo de ser el caso.

S&P	PI S&P	Modelo	PI Modelo
AAA	0%	R1	0.15%
AA+	0%		
AA+	0.02%		
AA	0.04%		
AA-	0.07%		
A+	0.09%		
A-	0.08%		
BBB+	0.16%	R2	0.45%
BBB	0.23%		
BBB-	0.38%	R3	0.79%
BB+	0.55%		
BB	0.80%	R4	2.41%
BB-	1.30%		
B+	2.60%	R5	4.20%
B	5.88%		
B-	9.12%	R6	12.81%
CCC+			
CCC	27.39%	R7	39.05%
CCC-			
D			

Tabla 3.1 Comparación de la PI contra la de agencia calificadorora

#### 4.2 Comparación del modelo contra el método regulatorio

Este modelo utiliza el mismo enfoque (un punto determinado en el tiempo) al darle mayor peso a las características propias del negocio, aunque las variables a las que les asigna mayor peso son un tanto diferentes a las ocupadas en el modelo, ya que casi no toma en cuenta la situación financiera que atraviesa en ese momento la empresa a calificar, ni la capacidad de pago, basándose principalmente en su historial de pagos y en cuestión de la probabilidad de incumplimiento que se puede observar en la tabla 3.2 el primer nivel de riesgo no hace mucha diferencia entre una buena empresa y una muy buena empresa al tener su primera PI un valor muy cercano al 1%.



Indicador	PI	Modelo Regulatorio	PI
R1	0.1%	A-1	0.9%
R2	0.5%	A-2	1.5%
R3	0.8%	B-1	2.0%
R4	2.4%	B-2	2.5%
R5	4.2%	B-3	5.0%
R6	12.8%	C-1	10.0%
R7	39.1%	C-2	15.5%
R8	65.0%	D	45.0%
R9	100%	E	100.0%

Tabla 3.2 Comparación de la PI contra el Modelo Regulatorio



## Capítulo 5 Resultados

Para la evaluación de los resultados que arroja el modelo de calificación crediticia se consideraron 4 empresas nacionales muy conocidas en México que han atravesado ciertos problemas en épocas recientes, se toma un horizonte de tiempo de 10 años evaluándolas trimestralmente y comparándolas con base en la probabilidad de incumplimiento del modelo regulatorio y las agencias calificadoras.

Las cuatro empresas elegidas al calificarse bajo el Método Regulatorio al tener ventas mayores a 14 millones de UDIS en todos los periodos analizados se calificaron bajo el Anexo 22 de la Circular Única de Bancos utilizando la modalidad de Grandes Corporativos al presentar todas ellas ventas mayores a 216 millones de UDIS.

La primera empresa a calificar fue CEMEX, la cual se centra en la industria de construcción teniendo como principales productos cemento, concreto premezclado y materiales de construcción, su sede se encuentra en Monterrey, Nuevo León, México y fue fundada en el año de 1906.

En las dos siguientes figuras se puede observar que la empresa CEMEX al ser calificada bajo el modelo interno, en tres periodos no consecutivos se alcanzó un nivel de riesgo mínimo de R6, el cual tiene una probabilidad de incumplimiento asociada de 12.8%. En la figura 4.1 se puede observar que la empresa al ser calificada bajo el Modelo Regulatorio, el nivel más bajo de riesgo alcanzado fue el de B-1 que tiene una probabilidad de incumplimiento asociada de 2%. En la figura 4.2 se observa que la calificación de Agencia mínima que obtuvo en el periodo analizado fue de B-, la cual tiene una probabilidad de incumplimiento asociada de 9.12%.

Se puede notar que la probabilidad de incumplimiento otorgada bajo el modelo de la CNBV es mucho menor al que obtuvo en los otros dos modelos, además de que la probabilidad de incumplimiento más baja se alcanzó en un punto distinto en el tiempo, así como que los modelos internos y el de agencia tienen una mayor similitud a la hora de degradar la calificación o aumentarla.

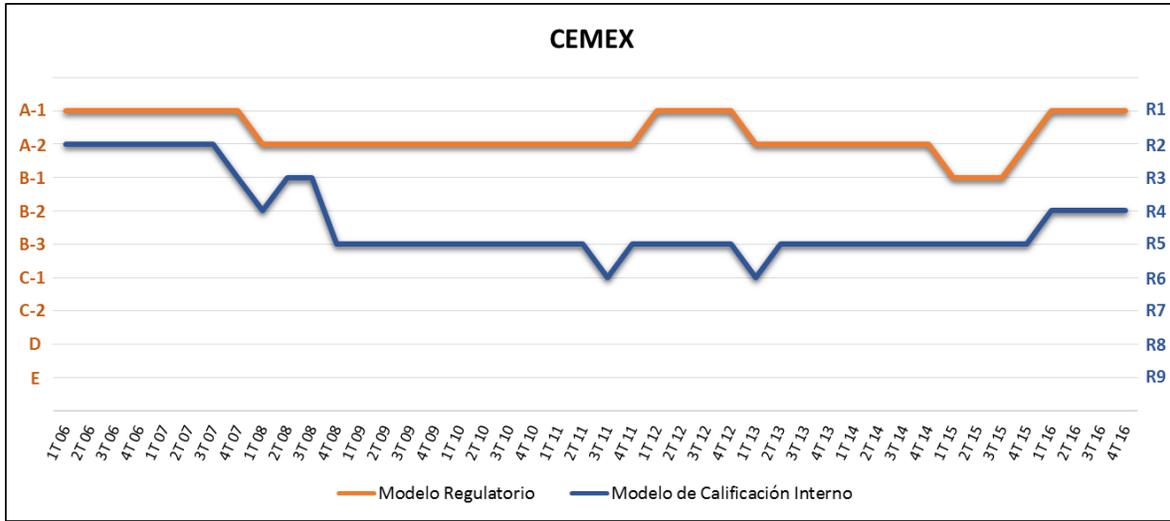


Figura 4.1 Gráfica comparativa del modelo regulatorio y el interno de la empresa CEMEX

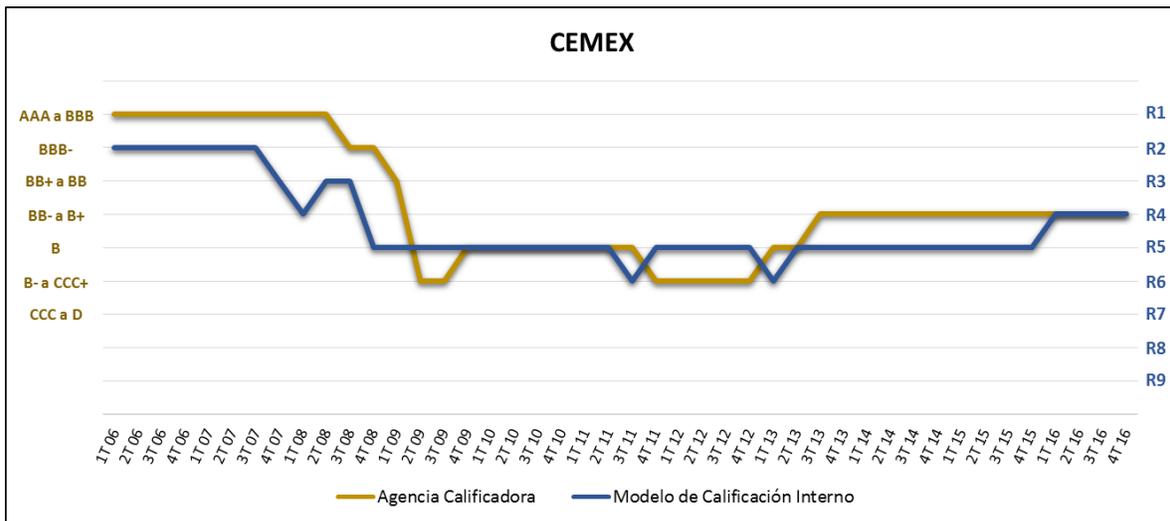


Figura 4.2 Gráfica comparativa de la calificación de agencia y el modelo interno de la empresa CEMEX



La segunda empresa a calificar fue Comercial Mexicana, la cual se centra en el comercio minorista bajo el formato de supermercado con sede en la Ciudad de México y la cual fue fundada en 1930. Actualmente ya no cotiza en la Bolsa Mexicana de Valores.

En las figuras 4.3 y 4.4 se puede observar que la empresa Comercial Mexicana al ser calificada bajo el modelo interno, alcanzó un nivel de riesgo mínimo de R5, el cual tiene una probabilidad de incumplimiento asociada de 4.2%. En la figura 4.3 se puede observar que la empresa al ser calificada bajo el Modelo Regulatorio, alcanzó su nivel más bajo de riesgo en un solo periodo y este fue de B-1 que tiene una probabilidad de incumplimiento asociada de 2%. En la figura 4.4 se observa que la calificación de Agencia mínima que obtuvo en el periodo analizado fue de CCC, la cual tiene una probabilidad de incumplimiento asociada de 27.39%.

En este caso se puede notar que la probabilidad de incumplimiento otorgada bajo el modelo de la CNBV es más parecida a la obtenida bajo el modelo interno y la de la agencia calificadora es mucho mayor a la que obtuvo en los otros dos modelos, en esta empresa se puede observar que hay una mayor similitud en el movimiento de las calificaciones.

En el periodo comprendido entre el tercer trimestre de 2009 y el segundo trimestre de 2011, la calificación de agencia se ve seriamente afectada debido a que ya se encontraba en reestructura y faltaba que cumpliera con las obligaciones derivadas de su reestructura.

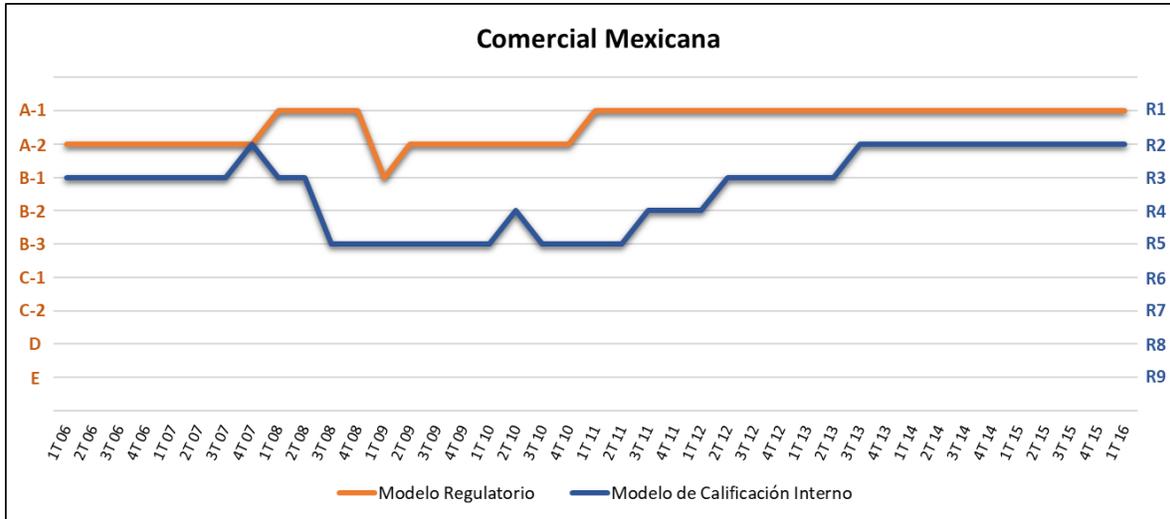


Figura 4.3 Gráfica comparativa del modelo regulatorio y el interno de la empresa Comercial Mexicana

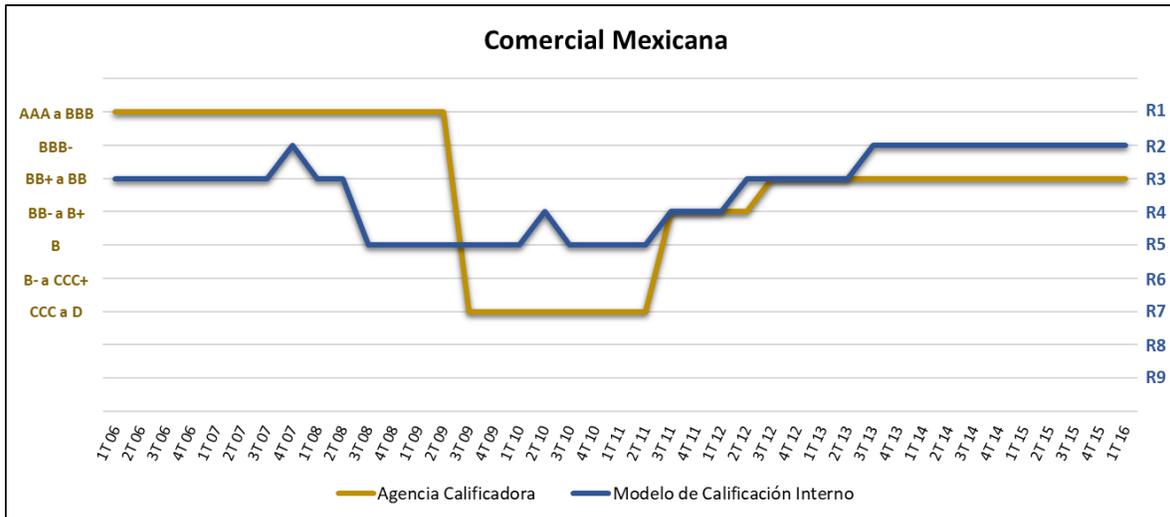


Figura 4.4 Gráfica comparativa de la calificación de agencia y el modelo interno de la empresa Comercial Mexicana

La tercera empresa en ser calificada fue Casas GEO, la cual se centra en construcción bajo los formatos de construcción de casas, su venta e intermediación de créditos, su sede se encuentra en la Ciudad de México y fue fundada en 1990.

En las figuras 4.5 y 4.6 se puede observar que la empresa Casas GEO al ser calificada bajo el modelo interno, alcanzó un nivel de riesgo mínimo de R7, el cual tiene una probabilidad de incumplimiento asociada de 39.1%. En la figura 4.5 se puede observar que la empresa al ser calificada bajo el Modelo Regulatorio, alcanzó su nivel más bajo de riesgo solamente en el



último periodo y este fue de B-1 que tiene una probabilidad de incumplimiento asociada de 2%. En la figura 4.6 se observa que la calificación de Agencia mínima que obtuvo en el periodo analizado fue de D, la cual tiene una probabilidad de incumplimiento asociada de 27.39%.

En este caso se puede notar que la probabilidad de incumplimiento otorgada bajo el modelo de la CNBV dista bastante de las obtenidas bajo los otros dos modelos, se puede observar como las calificaciones otorgadas por la agencia calificadora y por el modelo interno degradan la calificación de la empresa a partir del año 2013 para mantenerla en una calificación que indica un riesgo alto de incumplimiento mientras que en el caso del modelo regulatorio no se nota dicho cambio.

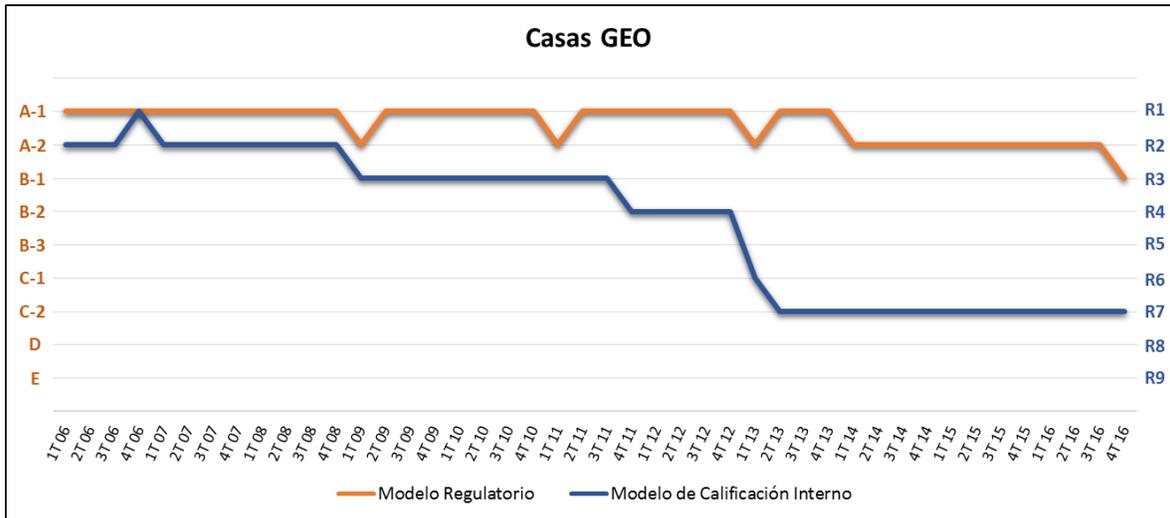


Figura 4.5 Gráfica comparativa del modelo regulatorio y el interno de la empresa Casas GEO

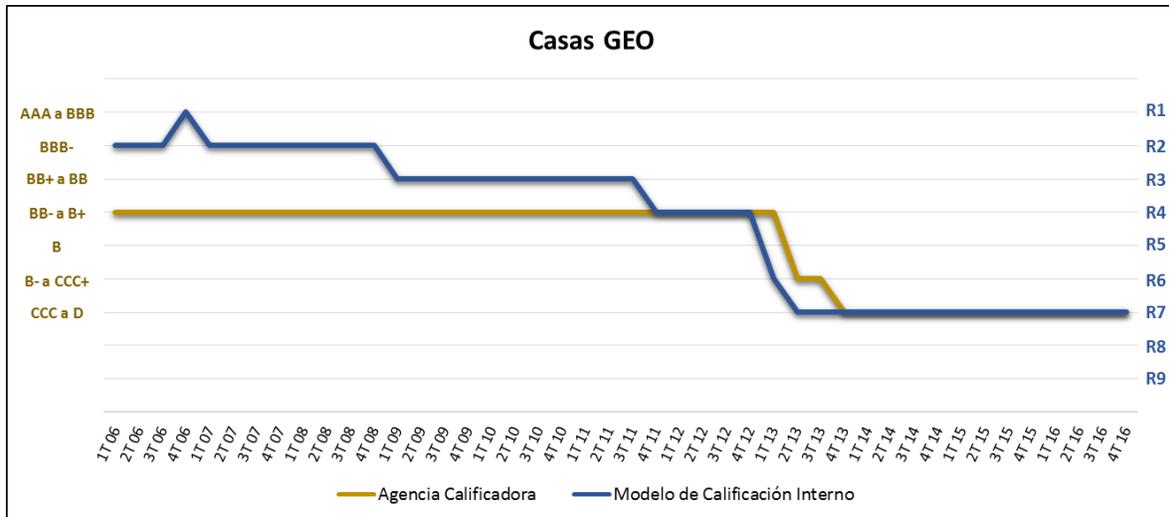


Figura 4.6 Gráfica comparativa de la calificación de agencia y el modelo interno de la empresa Casas GEO

La cuarta y última empresa en ser calificada fue ICA, esta empresa se centra en construcción en formato de infraestructura con sede en la Ciudad de México y fue fundada en el año de 1947.

En las últimas dos figuras se puede observar que la empresa ICA al ser calificada bajo el modelo interno, alcanzó un nivel de riesgo mínimo de R7, el cual tiene una probabilidad de incumplimiento asociada de 39.1%. En la figura 4.7 se puede observar que la empresa al ser calificada bajo el Modelo Regulatorio, alcanzó su nivel más bajo de riesgo en el último año y este fue de B-1 que tiene una probabilidad de incumplimiento asociada de 2%. En la figura 4.8 se observa que la calificación de Agencia mínima que obtuvo en el periodo analizado fue de CCC, la cual tiene una probabilidad de incumplimiento asociada de 27.39%.

En este caso se puede notar que la probabilidad de incumplimiento otorgada bajo el modelo de la CNBV dista bastante de las obtenidas al ser la empresa calificada bajo los otros dos modelos, se puede observar como las calificaciones otorgadas por la agencia calificadora y por el modelo interno están casi superpuestas casi todo el tiempo y como la empresa gradualmente ha ido degradando su calificación lo cual no se hace tan notorio en el modelo regulatorio.

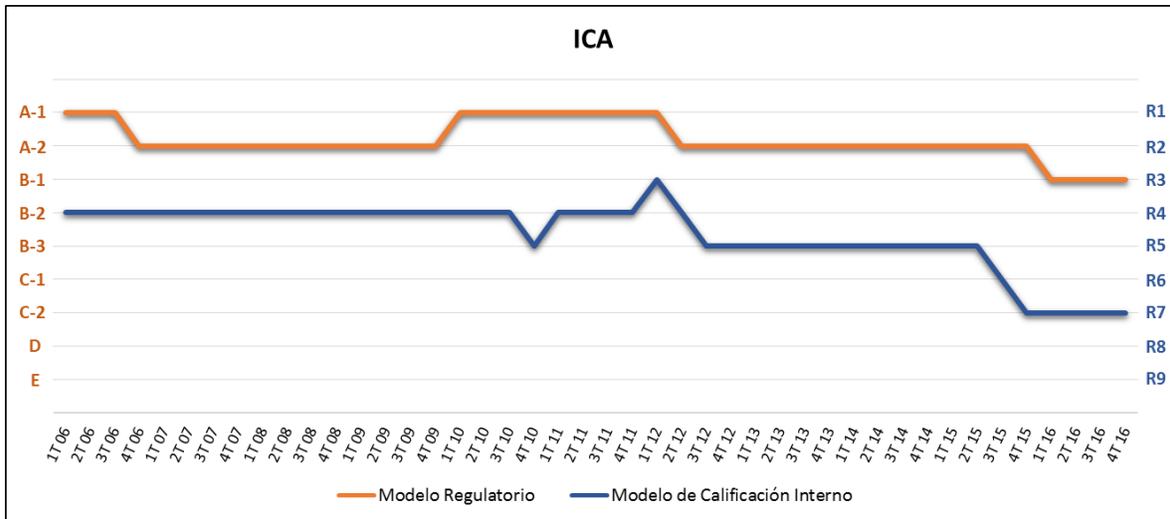


Figura 4.7 Gráfica comparativa del modelo regulatorio y el interno de la empresa ICA

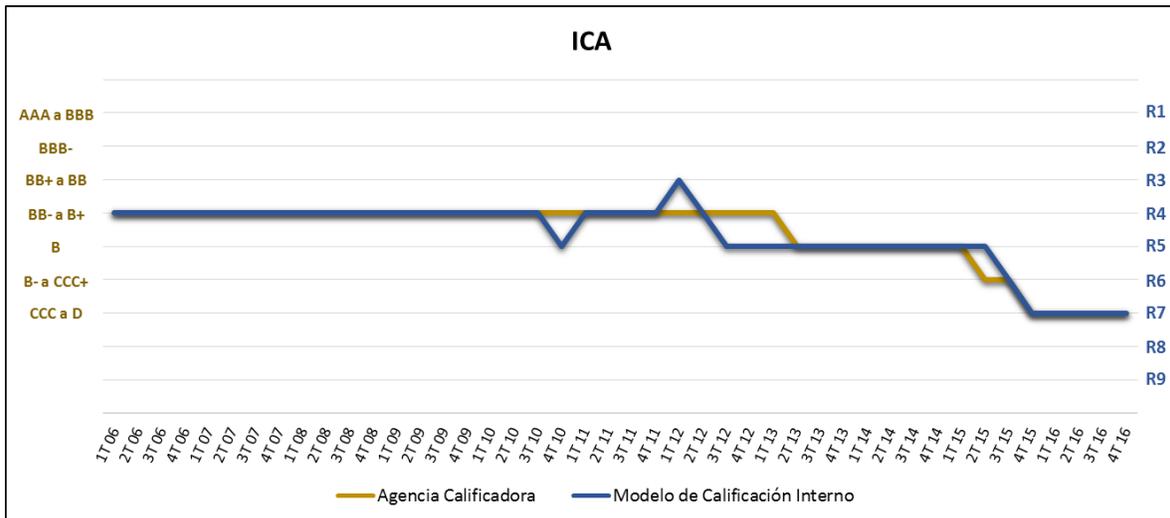


Figura 4.8 Gráfica comparativa de la calificación de agencia y el modelo interno de la empresa ICA

Como se puede apreciar, el modelo recoge los cambios en el nivel de riesgo de las empresas al igual que el modelo regulatorio, aunque se puede observar que por lo general las caídas de calificación son mayores ya que las empresas nunca dejaron de cumplir con sus pagos, aunque financieramente no se encontraran en su mejor momento lo que hace que la calificación regulatoria no los califique tan bajo. Mientras que la compararse con las calificaciones otorgadas por las agencias calificadoras se observa un parecido mayor y en algunos casos que el modelo interno actúa más rápidamente que el de las agencias.



## Capítulo 6 Conclusiones

En el trabajo se buscó actualizar el modelo de calificación crediticia derivado de los problemas observados al utilizar el modelo regulatorio para las etapas de *pricing* y seguimiento, con base en un análisis cuantitativo y cualitativo que le otorga mayor peso a la situación financiera en lugar de a la experiencia y voluntad de pago, complementándose con indicadores del sector, experiencia de pago, así como garantías ofrecidas.

En los resultados obtenidos de calificar a algunas empresas de gran tamaño e importancia en el país que han o todavía cotizan en la Bolsa Mexicana de Valores, y que en años anteriores han enfrentado distintos contratiempos se puede observar en cada una de las empresas evaluadas que la probabilidad de incumplimiento al utilizar el modelo regulatorio obtienen una calificación mínima de B-1 indicando una probabilidad de incumplimiento de 2% , dato que se usa para constituir las reservas crediticias regulatorias bajo la metodología de pérdida esperada, cifra que sería insuficiente en caso de que alguna de estas empresas cayera en estado de incumplimiento.

Esta calificación al otorgarle mayor peso a la experiencia y voluntad de pago no toma suficientemente en cuenta la situación financiera que la empresa atraviesa en dicho momento pudiendo encontrarse alguna empresa a punto de quebrar o estar en estado de disolución y si se encuentra al corriente de sus pagos o incumplió muy poco tiempo antes, la institución financiera puede tomar o encontrarse en un riesgo mucho mayor al estimado o al que tiene cubierto en dicho momento creando una pérdida no esperada de un monto razonable que irá contra el capital de la institución.

En la comparación hecha en el capítulo anterior se puede observar claramente como el método aprobado por la CNBV no mide satisfactoriamente el riesgo de un incumplimiento lo que podría ocasionar tomar un mayor riesgo de lo que la institución pueda manejar.



En cambio al compararse con la calificación otorgada por una agencia calificadora se puede observar una correlación positiva en cuanto a la medición del riesgo, pudiéndose también observar que la metodología propuesta responde más rápido que el modelo de las agencias a cambios en la situación financiera de las empresas.

Con base en esto se propone la incorporación del método interno antes expuesto para una mejor clasificación de los créditos en la fase de admisión y seguimiento de los mismos, esto sin ser tomada como una metodología interna que tenga que ser sancionada por la CNBV y tener una mejor.

Se propone que el método interno sea revisado con una periodicidad anual, validando que si alguna empresa llega a caer en incumplimiento, el modelo lo haya pronosticado correctamente y que la pérdida no esperada no afecte a la institución.



## Bibliografía

- 1) Venegas Martínez, F. (2011). *Riesgos financieros y económicos Productos derivados y decisiones económicas bajo incertidumbre* (Segunda Edición). México
- 2) Altman, E. I. (1968) .*Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy* (Journal of Finance). Estados Unidos
- 3) De Lara Haro, A. (2008) *Medición y Control de Riesgos Financieros* (Tercera Edición). México.
- 4) Schreiner, M. (2002) *Ventajas y desventajas del scoring estadístico para las microfinanzas* (Washington University)
- 5) Makowski, P. (1985) *Credit Scoring Branches Out: Decision Tree-Recent Technology* (Credit World 75)
- 6) Scoring. Recuperado de <https://www.bbva.com/es/scoring-programa-aprueba-credito/>
- 7) Disposiciones de carácter general aplicables a las instituciones de crédito (Circular Única de Bancos). Recuperado de:  
<http://www.cnbv.gob.mx/Paginas/Normatividad.aspx>  
Anexo 21 Recuperado de:  
<https://www.cnbv.gob.mx/Anexos/Anexo%2021%20CUB.pdf>  
Anexo 22 Recuperado de:  
<https://www.cnbv.gob.mx/Anexos/Anexo%2022%20CUB.pdf>
- 8) Apuntes de clase estadísticos del Dr. Juan Miguel Marín Diazaraque.  
Recuperados de:  
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/AMult/tema6am.pdf>  
Y  
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/AMult/tema3am.pdf>
- 9) Historia del Comité de Basilea. Recuperado de;  
<https://www.bis.org/bcbs/history.htm?m=3%7C14%7C573%7C76>



- 10) Basel Committee on Banking Supervision (2005) *Studies on the Validation of Internal Rating Systems*. (Working Paper No. 14.)
- 11) Beaver, H. W. (1966) *Financial Ratios as Predictors of Failure* (Journal of Accounting Research, vol.4, supplement)
- 12) Gutiérrez Girault, M.A. (2007) *Modelos de Credit Scoring –Qué, Cómo, Cuándo y Para Qué-*
- 13) Amat, O. (s.f.) *Scoring y Rating. Cómo se elaboran e interpretan*. Recuperado de: [http://www.supercontable.com/envios/articulos/BOLETIN\\_TU-ASESOR\\_02\\_2014\\_Articulo\\_5.htm](http://www.supercontable.com/envios/articulos/BOLETIN_TU-ASESOR_02_2014_Articulo_5.htm)
- 14) Montaña Moreno, J.J. (2002) *Redes Neuronales Artificiales Aplicadas al Análisis de Datos* (Tesis Doctoral)
- 15) Restrepo, A.F. y Orozco, S.A. (2015) *Modelo Predictivo de Quiebra Para Bancos Estadounidenses*
- 16) Silva Abad, M. (2016) *Valuación actuarial de un portafolio de tarjeta de crédito* (Tesis).
- 17) Fitch Ratings (2017) *Metodología de Calificación de Finanzas Corporativas*.
- 18) Morales Castro A. y Morales Castro J.A. (2014) *Crédito y cobranza*.
- 19) García Sánchez M. y Sánchez Barradas C. (2005) *Riesgo de Crédito en México: aplicación del modelo CreditMetrics*



## Anexos

### **Anexo 1.** Definiciones de la Circular Única de Bancos (CUB)

XXIX. Cartera Crediticia o Cartera de Crédito:

c) Comercial: a los créditos directos o contingentes, incluyendo créditos puente denominados en moneda nacional, extranjera, en UDIs, o en VSM, así como los intereses que generen, otorgados a personas morales o personas físicas con actividad empresarial y destinados a su giro comercial o financiero; incluyendo los otorgados a entidades financieras distintos de los de préstamos interbancarios menores a 3 días hábiles; las operaciones de factoraje, Descuento y Operaciones de Cesión de Derechos de Crédito; operaciones de arrendamiento financiero que sean celebradas con dichas personas morales o físicas; los créditos otorgados a fiduciarios que actúen al amparo de fideicomisos y los esquemas de crédito comúnmente conocidos como “estructurados”. Asimismo, quedarán comprendidos los créditos concedidos a entidades federativas, municipios y sus organismos descentralizados, cuando sean objeto de calificación de conformidad con las disposiciones aplicables.

La Cartera Crediticia estará sujeta a Calificación sin incluir aquellos créditos a cargo del Gobierno Federal o con garantía expresa de la Federación, registrados ante la Unidad de Crédito Público de la Secretaría de Hacienda y Crédito Público, del IPAB o del Banco de México.

LXXXVI. Instituciones: a las instituciones de crédito a que se refiere el artículo 2 de la Ley de Instituciones de Crédito.

LXXXVII. Instituciones Calificadoras: a las Instituciones Calificadoras de Valores incluidas en el Anexo 1-B de estas Disposiciones. También se considerará como Instituciones Calificadoras a aquellas que, atendiendo a los criterios contenidos en las presentes



Disposiciones dé a conocer la Comisión en la red electrónica mundial denominada Internet en el sitio <http://www.cnbv.gob.mx>.

**Anexo 2.** Medida de Adecuación Muestral Kaiser-Meyer-Olkin (KMO)

Es una medida de adecuación maestral ocupada en el análisis factorial y en el análisis de componentes principales la cual compara los coeficientes de correlación parciales (un indicador del grado de relación entre dos variables, eliminando la influencia del resto de las variables) con el coeficiente de correlación de Pearson, su fórmula es la siguiente:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} rp_{ij}^2} \quad 0 \leq KMO \leq 1$$

Donde:  $r_{ij}$  es el coeficiente de correlación de Pearson entre la variable  $i$  y la variable  $j$ .

$rp_{ij}$  es el coeficiente de correlación parcial entre la variable  $i$  y la variable  $j$ .

En esta medida cuanto más cercano a 1 sea el valor obtenido, la relación entre las variables es mayor, mientras que si el valor es menor a 0.5, la relación entre las variables es muy pequeña y no se puede realizar el análisis.

Se utiliza el siguiente criterio para determinar si la información nos permite continuar con el análisis de componentes principales o análisis factorial.

KMO	Criterio
$\geq 0.75$	Bien
$\geq 0.5$	Aceptable
$< 0.5$	Insuficiente



## *Glosario*

**CUB:** Se refiere a las Disposiciones De Carácter General Aplicables a las Instituciones de Crédito, comúnmente conocida como Circular Única de Bancos la cual es expedida por la Comisión Nacional Bancaria y de Valores.

**Dicotómica o variable dicotómica:** Variable que únicamente puede tomar dos posibles valores, generalmente, opuestos entre sí.

**EBIT:** Proveniente del idioma inglés Earnings before Interest and Taxes, cuya traducción sería Ingresos antes de intereses e impuestos, este indicador nos proporciona el resultado de una empresa sin tener en cuenta los impuestos o costes financieros.

**EBITDA:** Proveniente del idioma inglés Earnings before Interest, Taxes, Depreciation and Amortization, cuya traducción sería Ingresos antes de intereses, impuestos, depreciaciones y amortizaciones, este indicador nos proporciona el beneficio bruto calculado antes de la deducibilidad de gastos financieros.

**Eigenvalor:** Se refiere a un valor escalar que mantiene la dirección del vector al ser transformado, también conocido como valor propio, autovalor o valor característico.

**Eigenvector:** Se refiere a un vector no nulo que al multiplicarse por un escalar conocido como eigenvalor mantiene su dirección, también conocido como vector propio o vector característico.

**Ortogonal:** Un vector en un ángulo de  $90^\circ$ , de tal forma que son perpendiculares entre sí.

**Percéptron:** Unidad de interferencia en forma de discriminador lineal, a partir del cual se desarrolla un algoritmo.

**Pricing:** Proveniente del idioma inglés, cuya traducción podría ser poner precio, y se refiere a la actividad realizada para asignar una tasa de interés a cada financiamiento en base a las especificaciones de este.

**Scoring:** Se refiere al sistema de evaluación que se da de forma automática de solución de operaciones de crédito de préstamo.