

Modelos Para Calcular la Pérdida Esperada A Través Del Deterioro Crediticio

Samuel Alberto Gómez Meza
Escuela de Matemáticas y Ciencias de la Computación
Universidad Nacional Autónoma de Honduras
Supervisor: Doctor Fredy Antonio Vides Romero
16 de agosto de 2021

ÍNDICE

I.	Introducción	2
II.	Preliminares	2
III.	Planteamiento del Problema	3
III-A.	Cadenas de Markov	3
III-B.	Regresión Logística	3
III-C.	Valor en Riesgo	4
IV.	Resultados	4
IV-A.	Contraste de Hipótesis de Nulidad de los Parámetros (Test de Wald)	5
IV-B.	Parámetro M	6
IV-C.	Parámetro Loss Given Default	7
IV-D.	Matriz de Transición	7
IV-E.	Resultados Centrales	8
IV-E1.	Pérdida Esperada	8
IV-E2.	Valor en Riesgo (VaR)	8
IV-E3.	Pérdida Inesperada (PI)	8
IV-E4.	Resultados Matriz de Transición (Base de Datos Año 2019)	9
IV-E5.	Resultados Matriz de Transición (Base de Datos Año 2020)	9
V.	Conclusiones	10
VI.	Autocrítica	11
VII.	Disponibilidad de los Datos	11
VIII.	Agradecimientos	11
	Referencias	11

ÍNDICE DE FIGURAS

1.	Variables en la Ecuación (Test de Wald hacia Atrás)	6
----	---	---

ÍNDICE DE CUADROS

I.	Lista de Variables en la Base de Datos	6
II.	Resumen del Modelo Generado en R	6
III.	Cálculo del Parámetro M	6
IV.	Factor de Recuperación Efectiva del Valor de la Garantía	7
V.	Convenciones	7
VI.	Resultados Centrales Pérdida Esperada	8
VII.	Resultados Centrales Probabilidad Máxima de Incumplimiento	8
VIII.	Resultados Centrales Valor en Riesgo	8
IX.	Resultados Centrales Pérdida Inesperada	8
X.	Indicadores de Riesgo Financiero (Total del Portafolio)	9

XI.	Resultados Centrales Matriz de Transición (Momento Inicial - Momento Final (Año 2019))	10
XII.	Resultados Centrales Matriz de Transición (Momento Inicial - Mora Máxima (Año 2019))	10
XIII.	Resultados Centrales Matriz de Transición (Momento Inicial - Momento Final (Año 2020))	10
XIV.	Resultados Centrales Matriz de Transición (Momento Inicial - Mora Máxima (Año 2020))	10

Modelos Para Calcular la Pérdida Esperada A Través Del Deterioro Crediticio

DEDICATORIA

La presente Tesis de Investigación esta dedicada a la memoria de mi hermano Alexis Quintin Gómez Meza, quien siempre me animó, apoyó y creyó en mí durante estos años de estudio. Vivió su vida, ayudando tanto a familiares como a extraños necesitados, me enseñó el verdadero significado y importancia de la amistad, y sobre todo me enseñó que sin importar lo duro que fuera la vida siempre hay uno y mil motivos más para seguir adelante.

Resumen—En este artículo se busca analizar las metodologías utilizadas para medir el Riesgo de Crédito alusivo al incumplimiento por parte del cliente hacia una Institución Financiera. El análisis de Riesgo Crediticio, se realiza tanto nivel Individual (Micro) como a nivel de Portafolio de Créditos (Macro). Utilizando para el cálculo de la Probabilidad de Incumplimiento a nivel Micro el Modelo de Regresión Logística Binaria, y para cálculo de la probabilidad de incumplimiento a nivel Macro el Modelo de Cadenas de Markov (Matrices de Transición). Una vez estimada la Probabilidad de Incumplimiento a nivel Micro, se realiza el cálculo de los Indicadores de Riesgo Financiero (Pérdida Esperada, Valor en Riesgo, Pérdida Inesperada) para dar una interpretación simultanea de los mimos.

I. INTRODUCCIÓN

Los Modelos de Riesgo de Crédito proporcionan estimaciones del Riesgo, tales como la Pérdida Esperada, Valor en Riesgo y la Pérdida Inesperada [1]; el enfoque de medición del Riesgo crediticio se acompaña con la presentación de los modelos matemáticos y estadísticos, los requisitos de información, los procesos de estimación estadística y econométrica, interpretación de los resultados y consecuencias en la estrategia de la Institución Financiera.

Pero la mayoría de autoridades financieras desconocen la importancia de la gestión de riesgos desde un enfoque técnico y estructurado, sobre el cual se deben tomar decisiones que lleven al balance a la rentabilidad con el riesgo. La esencia de la actividad de las Instituciones Financieras y Tesorerías de grandes firmas es la toma de riesgos [1], por lo que el riesgo es ineludible en sus operaciones. En este sentido para poder beneficiarse del riesgo una vez identificado, es medirlo y de ser posible, predecirlo.

II. PRELIMINARES

Podemos definir el Riesgo Financiero como el grado de incertidumbre ligado al rendimiento o coste futuro de una operación financiera o bien como aquella condición en la cual existe una posibilidad de desviarse del resultado esperado o

deseado. [2]

Los riesgos a los que enfrenta una organización pueden ser de distinta naturaleza, entre los Riesgos Financieros más comunes tenemos: [2]

- Riesgo Crediticio
- Riesgo de Mercado
- Riesgo de Liquidez

El riesgo se puede cubrir de cierta manera y sobre todo se puede mitigar siempre y cuando seamos capaces de:

- Conocer el riesgo al que nos exponemos.
- Medirlo: es decir valorando la probabilidad de que este produzca.
- Gestionarlo: actuando si este se produce.

La Pérdida Esperada se define: como el costo que enfrenta la Institución Financiera por el Riesgo de Crédito que le representa el cliente; es decir el monto de capital que podría perder una institución como resultado del incumplimiento por parte del cliente [3], en algunos países se le conoce como provisión de crédito y en otros como previsión. La Pérdida Inesperada o capital en riesgo, representa la cantidad que la institución financiera puede perder por arriba de la pérdida esperada. [4]

La metodología propuesta por Basilea II para el cálculo de la Pérdida Esperada surge a partir de observar el Riesgo de Crédito que contempla dos criterios para su medición (Características del cliente y Características de la línea de crédito), la Probabilidad de Incumplimiento que es la posibilidad de que un cliente no pueda cubrir sus obligaciones de pago en un periodo determinado, el Saldo Actual de la Operación es el monto contable que el cliente adeuda al momento en que se presenta el incumplimiento, incluye capital e intereses. Por otra parte la Severidad de la Pérdida (LGD - loss - Given - Default) se determina por el tipo y el valor de las garantías ligadas al crédito evaluado. [4]

El Modelo de Pérdida Esperada se obtiene al multiplicar estos parámetros individuales quedando de la siguiente forma:

$$PE = EC \cdot LGD \cdot PD \cdot M \quad (1)$$

Donde:

PE: Pérdida Esperada.

EC: Saldo actual de la operación; monto expuesto; exposición de crédito.

LGD: Loss - Given - Default o Proporción de crédito que esta al descubierto porque el valor de la garantía es insuficiente para cubrir el monto de la exposición.

PD: Probabilidad de Incumplimiento del cliente.

M: Parámetro de Maduración.

De los factores necesarios para el cálculo de la Pérdida Esperada, la Probabilidad de Incumplimiento (PD) requiere el uso de técnicas estadísticas para estimar la probabilidad de que un cliente suspenda el servicio de su operación de crédito, las cuales pueden ser:

- Cadenas de Markov.
- Regresión Logística Binaria.
- Árbol de decisión.
- Redes Neuronales o K Vecinos más Cercanos.

Cualquiera de estas técnicas demandan el uso de información personal, demográfica, laboral, económica, financiera, social, de cada uno de los clientes. El modelo calcula tres indicadores básicos de la exposición al Riesgo de Crédito:

- La Pérdida Esperada.
- El Valor en Riesgo (VaR).
- La Pérdida Inesperada (Capital en Riesgo).

La explicación financiera del Valor en Riesgo (VaR) responde a la posibilidad de que dadas las condiciones del momento, la Pérdida Esperada pudiese variar entre cero (Recuperación Plena del Crédito) hasta un máximo sujeto a un nivel de confianza dado. El calculo del Valor en Riesgo sigue de la siguiente forma:

$$VaR = EC \cdot LGD \cdot Pmax \quad (2)$$

Donde:

VaR: Valor en Riesgo.

EC: Exposición de Crédito.

LGD: Proporción del crédito al descubierto porque el valor de la garantía es insuficiente para su cobertura total.

Pmax: Máxima Probabilidad de Incumplimiento.

Para el cálculo de la Pérdida Inesperada que dicho de otra forma; es el valor que una institución no espera perder, se calcula mediante la siguiente formula:

$$PI = VaR - PE \quad (3)$$

III. PLANTEAMIENTO DEL PROBLEMA

El objetivo principal es analizar las metodologías utilizadas para medir el riesgo de crédito derivado del incumplimiento por parte del cliente, el análisis que se lleve a cabo por parte de cada institución financiera debe contemplar el riesgo de crédito tanto en las operaciones individuales como al análisis de riesgo de todo el Portafolio, no obstante el análisis de riesgo de crédito puede contemplarse bajo varios modelos estadísticos tanto a nivel individual como a nivel de Portafolio. [1]

Para el estudio posterior se utilizarán Cadenas de Markov para estimar el deterioro del total del Portafolio por tipo de operación, dichas operaciones serán segmentadas por tipo de crédito, los cuales son: créditos comerciales, vivienda, consumo y tarjetas. Por otra parte el análisis mediante regresión logística contempla la misma segmentación con la diferencia que se enfocará en análisis individuales de las operaciones de los clientes.

III-A. Cadenas de Markov

Los procesos de paso aleatorio son un caso particular de procesos estocásticos más generales que son las Cadenas de Markov. Una cadena es un proceso en tiempo discreto en el que una variable aleatoria X_n va cambiando con el paso del tiempo. Una de las propiedades de las Cadenas de Markov es que la probabilidad de que $X_n = j$ solo depende del estado anterior del sistema X_{n-1} . [5]

La matriz de transición es una herramienta fundamental para el cálculo de la Pérdida Esperada la cual se utilizará para determinar la probabilidad de que un crédito se deteriore en un periodo específico, comúnmente es conocida en la teoría como la Probabilidad de Incumplimiento.

Las Matrices de Transición representan un elemento importante en la estimación del Riesgo de Crédito para el sistema financiero, debido a que prevén las posibles pérdidas debido al deterioro que pudiera presentar una cartera en el futuro. [1] [3]

El procedimiento de las Matrices de Transición pueden platearse como:

- p_{ij} la probabilidad de que un deudor con calificación crediticia i en el tiempo t pueda migrar a otra calificación crediticia j en un tiempo $t + 1$ determinado.

Una vez definidas las probabilidades p_{ij} es posible construir la matriz de probabilidades de transición P con i filas y j columnas. Al ser una matriz de probabilidades satisface automáticamente las propiedades $p_{ij} \geq 0$, $\sum p_{ij} = 1$, $\forall i$.

Las probabilidades pueden estimarse de la siguiente manera:

$$P_{ij} = \frac{N_{ij}}{N_i}, \forall i, j \quad (4)$$

Donde:

N_{ij} : Total de créditos que pasaron de la calificación i a la calificación j en un periodo $t + 1$.

N_i : Total de créditos que estaban en la calificación i en el periodo t . [1] [3]

A partir de la matriz P se obtienen las diferentes matrices de transición para t periodos elevando la matriz P a la potencia t .

$$P_t = \begin{bmatrix} P_{00}^t & P_{01}^t \\ P_{10}^t & P_{11}^t \end{bmatrix} \quad (5)$$

Donde P_{ij}^t representa la probabilidad de pasar de la calificación i a la calificación j en t periodos.

III-B. Regresión Logística

Si una variable cualitativa consta de dos niveles y se codifica como 0 y 1, matemáticamente es posible ajustar un modelo de Regresión Lineal por Mínimos Cuadrados $\beta_0 + \beta_1 X$. El problema que presenta esta aproximación del modelo, es que al tratarse de una recta, para valores extremos

del predictor, se obtienen valores de Y menores que 0 y mayores que 1, lo que entra en contradicción con el hecho de que las probabilidades siempre están dentro del rango $[0, 1]$.

Para evitar este problema, la Regresión Logística transforma el valor devuelto por la Regresión Lineal $\beta_0 + \beta_1 X$ empleando una función cuyo resultado está siempre comprendido en el rango $[0, 1]$. Existen varias funciones que cumplen esta descripción, una de las utilizadas es la Función Logística (También conocida como Función Sigmoide): [6]

$$\sigma(x) = \frac{1}{1 + e^x}$$

El Modelo de Regresión Logística para la estimación de la Probabilidad de Incumplimiento de las carteras de crédito de las Instituciones Financieras presenta varias ventajas con respecto al Modelo utilizando Cadenas de Markov (Matrices de Transición); ya que este modelo permite además de obtener estimaciones consistentes de la Probabilidad de Incumplimiento, identificar los factores de riesgo que determinan dichas probabilidades y el peso o influencia de éstos sobre las mismas. [1]

Una vez que la Institución Financiera ha definido el punto crítico para la evaluación del Riesgo de Crédito, se requiere asignar un código binario para cada operación, de esta manera poder identificarlo como un cliente que cumple con sus obligaciones o un cliente que incumple con sus obligaciones crediticias. Para el caso de la Regresión Logística Binaria el código asignado para cada operación será la siguiente: cero si el cliente presenta una mora menor de 90 días y uno si el cliente presenta una mora mayor o igual de 90 días, dicho margen de tiempo esta estipulado en las circulares de Basilea II.

$y_i = 0$ Si el Cliente está en un estado de cumplimiento
 $y_i = 1$ Si el Cliente está en un estado de incumplimiento

La técnica estadística más tradicional y clásica para estimar una variable dicotómica como la y_i , es la Regresión Logística Binaria, la forma que tiene dicha Regresión es la siguiente: [6]

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad (6)$$

Donde:

y_i : Variable Dependiente con valor 0 ó 1.

X_i : Variables Independientes. (Demográficas, Laborales, Financieras, Económicas).

β_i : Parámetros de la Regresión; coeficientes que miden el impacto de variaciones de las variables X_i en el comportamiento del pago y_i .

μ_i : Terminio de error de la Regresión.

Una aplicación practica de la Regresión Logística Binaria es la estimación del margen de Pérdida Esperada que debería la institución financiera cobrar al cliente según su nivel de Probabilidad de Incumplimiento. [1] [6]

$$PD = \frac{e^y}{1 + e^y} \quad (7)$$

III-C. Valor en Riesgo

Desde el punto de vista estadístico, la Pérdida Esperada es una estimación estadística sujeta a los errores de muestreo y de la calidad y cantidad de información con que se cuenta. El Valor en Riesgo es un concepto asociado con la variabilidad de las Pérdidas Esperadas de cada una de las operaciones de crédito. Para encontrar la variabilidad es necesario calcular la desviación estándar de cada Pérdida Esperada por lo que necesitamos conocer previamente la varianza.

Para el cálculo del Valor en Riesgo (VaR) ecuación (2), necesitamos encontrar primero P_{max} mediante la siguiente formula:

$$P_{max} = N \left[\frac{1}{\sqrt{1-\rho}} \cdot G(PD) + \sqrt{\frac{\rho}{1-\rho}} \cdot G(0,999) \right] \quad (8)$$

Donde:

P_{max} : Probabilidad Máxima de Incumplimiento; Basilea no le designa ningún nombre específico; se calcula como una aproximación de la distribución de probabilidad Normal Estándar.

$N[\dots]$: Función de distribución de probabilidad Normal Estándar.

ρ : Coeficiente de riesgo sistémico o de contagio dentro de la cartera de crédito.

$G(\dots)$: Valor critico de la función de distribución Normal Estándar; valor inverso de la probabilidad.

Es fiable suponer que entre las Pérdidas Esperadas de cada operación de crédito, pueda existir algún grado de asociación que contagie la solvencia de una operación o varias; siendo esta condicionada por la solvencia de otra operación, por lo que el modelo avanzado de Basilea aproxima el riesgo de contagio (ρ) mediante el cálculo de las raíces características de la matriz de varianzas y covarianzas, estas raíces características son las siguientes: [9] [8]

1. $\frac{1}{1-\rho}$
2. $\frac{\rho}{1-\rho}$

Para la simplificación del cálculo del riesgo de contagio, el modelo avanzado de Basilea propone tres valores según el producto de crédito de las carteras de las personas físicas: [9]

- a). $\rho = 0,15$, Para carteras de vivienda.
- b). $\rho = 0,04$, Para créditos devolutivos.
- c). Para el resto de las carteras masivas, la siguiente fórmula:

$$\rho = 0,03 \cdot \frac{1 - e^{-35 \cdot PD}}{1 - e^{-35}} + 0,16 \cdot \frac{1 - (1 - e^{-35 \cdot PD})}{1 - e^{-35}}$$

IV. RESULTADOS

La información (Base de Datos) y los resultados generados utilizando los software estadísticos R y SPSS; utilizados para el estudio del cálculo de la Probabilidad de Incumplimiento

utilizando como modelo estadístico - matemático la Regresión Logística Binaria esta disponible en el repositorio de Github: <https://github.com/SM-Gomez>.

IV-A. Contraste de Hipótesis de Nulidad de los Parámetros (Test de Wald)

La Regresión Logística se distingue del modelo de Regresión Lineal Múltiple en el hecho de que en la Regresión Logística las variables no deben ser necesariamente cuantitativas ni tampoco cumplir supuestos de normalidad, pese a que la solución puede ser más estable si los predictores tienen una distribución normal multivariante. [10]

Los Supuestos que se deben cumplir para el uso de la Regresión Logística son los siguientes: [10]

- Las Variables Predictoras deben de ser categóricas o continuas.
- No se requiere Principio de Linealidad.
- Independencia del Error.
- No debe existir Multicolinealidad entre las Variables.

Por lo que la base de datos que se va utilizar para el cálculo de la Probabilidad de Incumplimiento mediante el uso de la Regresión Logística Binaria se describe en el cuadro I.

De estas variables no todas tienen un valor significativo relevante para ser utilizadas en el modelo de Regresión Logística Binaria, por lo que se utilizará el Test de Wald para seleccionar las variables que mejor se ajusten al modelo, para la estimación del Test se hará bajo el ambiente de SPSS con el fin de crear el modelo que mejor prediga nuestra variable dependiente (Estado del Crédito).

Para conocer el contraste de hipótesis de nulidad de los parámetros; el Test de Wald es uno de los más usados que se efectúa para cada una de las variables que intervienen en el modelo.

Para un coeficiente cualquiera β_i , se verifica (Para muestras suficientemente grandes) que bajo la hipótesis nula $H_0 : \beta_i = \beta_0$, el estadístico w definido por:

$$w = \frac{(\beta_i - \beta_0)^2}{Var(\beta_i)} \quad (9)$$

Se distribuye como una $\chi^2_{(1)}$ si la variable explicativa es cuantitativa y si la variable explicativa es de tipo categórica, se distribuye como una $\chi^2_{(c-1)}$, donde c es igual al número de categorías que toma la variable explicativa; para nuestro caso las variables explicativas categóricas constan de dos categorías por lo tanto $c = 2$, entonces $\chi^2_{(c-1)} = \chi^2_{2-1} = \chi^2_{(1)}$.

El estadístico w permite establecer qué variables son importantes para explicar la probabilidad del suceso ($y = 1$), mediante el contraste de hipótesis:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

La decisión se toma en base a la significancia del estadístico w o si el $p - \text{valor}$ es inferior a 0,05, se rechaza H_0 al nivel α del 5%, en caso contrario no se rechaza H_0 . [11] [12]

Aplicando el Test de Wald hacia atrás, veamos que esta pasando con las variables explicativas (vea la figura 1):

- Para las variables explicativas: Saldo de la Cuota, Tasa de Interés, Dependencia Ocupacional, se rechaza la hipótesis nula, por lo tanto dichas variables son significativas.
- Para las Variables Explicativas:

Variable	Wald	p – valor
Número de Cuotas	0.059	0.808
Monto del Crédito	0.000	0.983
Ingreso Mensual	2.533	0.112
Sexo	0.555	0.456

Entonces diremos que no se rechaza la hipótesis nula; por lo tanto dichas variables no son significativas.

Una desventaja al utilizar el test de Wald para conocer el contraste de hipótesis de nulidad de los parámetros, es que se puede cometer el error de tipo II durante se realiza la selección de variables; por este motivo se genera en el software estadístico R haciendo uso de las librerías: ggplot2, MASS, mlogit, dplyr, tidyr, readxl; un modelo de Regresión Logística Binaria usando las variables explicativas: Saldo de la Cuota, Tasa de Interés, Dependencia Ocupacional, Número de Cuotas, Monto del Crédito, Ingreso Mensual y Sexo, se crea dicho modelo con el objetivo de conocer el Resumen del mismo y ver que variables en realidad son las más significativas.(Ver Cuadro II)

Del Modelo creado en R al observar el resumen del mismo, se observa que todas las variables explicativas son significativas. Considerando los resultados obtenidos en I y II, las variables explicativas y significativas para el modelo de Regresión Logística Binaria que se utilizará para calcular la Probabilidad de Incumplimiento (PD) queda de la siguiente forma:

$$PD = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7}} \quad (10)$$

Variable	Tipo de Dato
Identificación del Crédito	Escala
Saldo Vigente	Escala
Días de Atraso	Escala
Estado del Crédito	Nominal
Actividad Económica	Nominal
Saldo de la Cuota	Escala
Permanencia en la Vivienda	Escala
Número de Cuotas	Escala
Monto del Crédito	Escala
Tasa de Interés	Escala
Ingreso Mensual	Escala
Sexo	Nominal
Frecuencia de Pago	Nominal
Dependencia Ocupacional	Nominal
Tipo de Cuota	Nominal
Profesión del Cliente	Nominal
Bienes Raíces	Nominal
Automóviles	Nominal

Cuadro I
LISTA DE VARIABLES EN LA BASE DE DATOS

Coefficiente	Estimado (β_i)	Significancia
Intercepto (β_0)	0.2117	-
Saldo de la Cuota (X_1)	-0.01016	0.001
Número de Cuotas (X_2)	-0.01341	0.02731
Monto del Crédito (X_3)	-0.00001178	0.00875
Tasa de Interés (X_4)	-0.1218	0.001
Ingreso Mensual (X_5)	0.00005421	0.00157
Sexo (X_6)	0.2821	0.01977
Dependencia Ocupacional (X_7)	0.7670	0.001

Cuadro II
RESUMEN DEL MODELO GENERADO EN R

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a						
Saldo_de_la_Cuota	-.010	.002	30.842	1	.000	.990
Numero_de_Cuotas	-.002	.009	.055	1	.815	.998
Monto_del_Credito	.000	.000	.000	1	.983	1.000
Tasa_de_Interes	-.055	.026	4.367	1	.037	.946
Ingreso_Mensual	.000	.000	2.362	1	.124	1.000
Sexo(1)	.144	.198	.530	1	.466	1.155
Dependencia_Ocupacion al(1)	-.847	.207	16.708	1	.000	.429
Constant	-.218	.797	.075	1	.784	.804
Step 2 ^a						
Saldo_de_la_Cuota	-.010	.001	55.777	1	.000	.990
Numero_de_Cuotas	-.002	.009	.059	1	.808	.998
Tasa_de_Interes	-.055	.025	4.811	1	.028	.946
Ingreso_Mensual	.000	.000	2.388	1	.122	1.000
Sexo(1)	.144	.198	.534	1	.465	1.155
Dependencia_Ocupacion al(1)	-.847	.207	16.781	1	.000	.429
Constant	-.222	.774	.083	1	.774	.801
Step 3 ^a						
Saldo_de_la_Cuota	-.010	.001	60.736	1	.000	.990
Tasa_de_Interes	-.053	.024	4.955	1	.026	.948
Ingreso_Mensual	.000	.000	2.602	1	.107	1.000
Sexo(1)	.147	.197	.555	1	.456	1.159
Dependencia_Ocupacion al(1)	-.857	.202	17.903	1	.000	.425
Constant	-.324	.651	.248	1	.619	.723
Step 4 ^a						
Saldo_de_la_Cuota	-.010	.001	60.353	1	.000	.990
Tasa_de_Interes	-.052	.024	4.816	1	.028	.949
Ingreso_Mensual	.000	.000	2.533	1	.112	1.000
Dependencia_Ocupacion al(1)	-.849	.202	17.619	1	.000	.428
Constant	-.258	.644	.161	1	.688	.772
Step 5 ^a						
Saldo_de_la_Cuota	-.010	.001	58.775	1	.000	.990
Tasa_de_Interes	-.054	.024	5.132	1	.023	.947
Dependencia_Ocupacion al(1)	-.876	.201	18.911	1	.000	.417
Constant	-.185	.645	.082	1	.775	.831

Figura 1. Variables en la Ecuación (Test de Wald hacia Atrás)

IV-B. Parámetro M

La incorporación del Parámetro de Plazo al Vencimiento (Parámetro M), pretende capturar en el cálculo de la Pérdida Esperada el concepto que mientras mayor sea el plazo al vencimiento, mayor será la Pérdida Esperada. [4] [9]

El Parámetro M puede asumir una de las siguientes posibilidades:

- Si la operación ya está vencida y continúa en el balance, entonces $M = 0$. Es importante tener presente que una operación vencida, ya sea por un día o varios, ya no representa una Pérdida Esperada.
- Su atraso en la cancelación ya no es un factor de riesgo de crédito, sino más bien debe ser atendido por la gestión de cobro.
- Si a la operación de crédito le resta para su vencimiento menos de un año, entonces $M = 1$.
- Si a la operación de crédito le resta entre un año y menos de tres años, entonces $M = 1.025$; es decir, la Pérdida Esperada aumentaría en un 2.5 %.
- Si a la operación de crédito le resta entre tres años y menos de cinco años, entonces $M = 1.05$; es decir, la Pérdida Esperada aumentaría en un 5 %.
- Si a la operación de crédito le resta cinco años o más, entonces $M = 1.075$; es decir, la Pérdida Esperada aumentaría en un 7.5 %.

En el cuadro III se muestra como se realiza el cálculo del Parámetro M, los resultados de realizar dicho cálculo a la base de datos con la que se está trabajando utilizando el software estadístico R se encuentran disponibles en el repositorio de Github.

Para realizar el cálculo necesitamos conocer la Fecha de Análisis (FA) y la Fecha de Vencimiento (FV), conociendo estas dos fechas procedemos a realizar la resta de la Fecha de Vencimiento menos la Fecha de Análisis esto se realiza con el fin de conocer los Días al Vencimiento (DaV), dicha cantidad de días será dividida entre un año comercial (365 días) para de esta manera obtener los Años al Vencimiento (AaV) y una vez obtenidos los años podemos calcular el Parámetro M.

FA	FV	DaV	AaV	M
5/31/2011	5/15/2012	350	0.96	1.000
5/31/2011	5/8/2015	1438	3.94	1.050
5/31/2011	4/15/2012	320	0.88	1.000
5/31/2011	10/15/2013	868	2.38	1.025
5/31/2011	2/10/2012	255	0.70	1.000

Cuadro III
CÁLCULO DEL PARÁMETRO M

IV-C. Parámetro Loss Given Default

La LGD, Pérdida Dado el Incumplimiento o Severidad de la Pérdida (en inglés Loss Given Default), mide aquella proporción del saldo vigente de la operación de crédito que estaría al descubierto por cuanto el valor de la garantía o colateral es insuficiente para cubrir la potencial pérdida en su totalidad.

Para realizar el cálculo de LGD se requiere conocer el Tipo de Garantía, el Valor de Mercado o Valor Razonable de la Garantía o Colateral, actualizado para una fecha lo más reciente posible y el Parámetro de Recuperación, el cual refleje el monto efectivo que la institución financiera esperaría recobrar del valor de la garantía, dicho parámetro puede variar desde cero hasta 100 por ciento. [13]

La formula para realizar el cálculo de LGD es la siguiente: [13]

$$LGD = 1 - f \quad (11)$$

Donde:

LGD : Severidad de la Pérdida

f : Tasa de Recuperación del crédito.

$$f = \frac{(LTV) \cdot (\text{Parámetro de Recuperación})}{EC} \quad (12)$$

Donde:

LTV : Valor de Mercado o Valor Razonable de la Garantía (Loan-to-Value).

EC : Exposición del Crédito o Saldo Vigente de la Operación.

El cálculo de LGD a la base de datos con la que se esta trabajando se realizaron en el software estadístico R los cuales están disponibles en el repositorio de Github [7], dicha base de datos presenta cuatro tipos de garantías y el factor de Recuperación asociado a dicha garantía son los estándares sugeridos por Basilea los cuales se describen en el siguiente cuadro:

Código	Tipo de Garantía	Factor de Recuperación
0	Sin Garantía	0.0 %
1	Back-to-Back	100.0 %
2	Personal	0.0 %
3	Hipotecaria Residencial	85.0 %

Cuadro IV

FACTOR DE RECUPERACIÓN EFECTIVA DEL VALOR DE LA GARANTÍA

IV-D. Matriz de Transición

Una Matriz de Transición es un arreglo de filas y columnas que se modifica, evoluciona y que es dinámico a través del tiempo. Es normal preguntarse que horizonte de tiempo se va a tomar como referencia, dicho horizonte de tiempo va a depender del problema al cual se va enfocar los resultados de esta metodología.

Las Matrices de Transición se construyen para segmentos de la población que son homogéneos, estas Matrices no

pretenden cubrir la totalidad de la cartera dado que se pueden encontrar elementos que no son totalmente homogéneos.

Aplicando esta metodología para calcular la Probabilidad de Incumplimiento (PD), el horizonte de tiempo mínimo requerido de es de 12 meses, por lo que las bases de datos con las que se trabajará constan de 12 vectores, la información que tienen dichos vectores corresponde a los Días de Mora tomados mensualmente para cada operación. Algunas Instituciones Financieras trabajan asignando una calificación a cada operación en lugar de Días de Mora.

Nos vamos a centrar en calcular la Probabilidad de Incumplimiento haciendo uso de los vectores de Días de Mora los cuales están comprendidos:

- Días de Mora Enero 2019 - Días de Mora Diciembre 2019.
- Días de Mora Febrero 2020 - Días de Mora Febrero 2021.

Dichos Días de Mora han sido identificados para un total de 5768 operaciones para el año 2019 y 1628 operaciones para el año 2020.

De esta manera se tiene un Horizonte de Tiempo de 12 meses y un Horizonte de Transición de 11 meses, estos horizontes son diferentes dado que se va a comparar el momento inicial (Días de Mora Enero) contra lo que le va sucediendo a cada una de las operaciones al transcurrir el tiempo, es decir tenemos un total de 11 meses para darle seguimiento a dichas operaciones.

Antes de aplicar el Modelo utilizando Matrices de Transición, se requiere establecer ciertos rangos o estados que son los que finalmente se van a comparar. Los estados se describen en el cuadro V.

Estado (Inicio/Fin)	Limite Inferior	Categoría
Al Día	0	000-000
1-30 Días	1	001-030
31-60 Días	31	031-060
61-90 Días	61	061-090
91-120 Días	91	091-120
121-150 Días	121	121-150
151-180 Días	151	151-180
181-210 Días	181	181-210
211-240 Días	211	211-240
241-270 Días	241	241-270
271-300 Días	271	271-300
301-330 Días	301	301-330
331-360 Días	331	331-360
Mayor a 360 Días	361	Mayor a 360

Cuadro V
CONVENCIONES

Estas convenciones se pueden modificar dependiendo de los resultados que desee obtener la Institución Financiera, para este caso vamos a trabajar con las convenciones propuestas en el cuadro V, las cuales están construidas en intervalos de 30 días, de tal manera tenemos un total de 14 estados, teniendo en cuenta que la mora que es mayor a 360 días queda clasificada como Mora Mayor a 360.

La Matriz de Transición se va a desarrollar utilizando una Tabla Dinámica en la cual se va a comparar los vectores Rango Inicial contra Rango Final, la información que contienen estos vectores son el total de créditos que empezaron en un estado (Rango Inicial) de los descritos en el cuadro V y al transcurrir el tiempo migraron hacia otro estado o se mantuvieron en el mismo (Rango Final). Las Tablas Dinámicas nos permiten determinar, comparar y analizar grandes sumas de información, sin tener que recurrir a conteos extensos o procesos estadísticos complejos o formulados. Dicha Tabla Dinámica se realizará en el software estadístico R haciendo una de las librerías devtools y rpivotTable.

Se van utilizar los estados descritos en el cuadro V para poder clasificar los días de mora en cada uno de los vectores de las bases de datos con las que se va a trabajar, dichas bases de Datos están disponibles en el repositorio de Github [7] al igual que los resultados obtenidos haciendo uso del software estadístico R.

IV-E. Resultados Centrales

Previamente se han realizado cálculos de Resultados Preliminares con el fin de facilitar el entendimiento que conlleva el cálculo de Resultados mas concretos o Centrales.

IV-E1. Pérdida Esperada: Utilizando la ecuación (10) se estimó la PD (Probabilidad de Incumplimiento), se realizaron los cálculos respectivos para el Parámetro M y finalmente haciendo uso de las ecuaciones (11) y (12) se realizaron los cálculos de LGD (Loss Given Default). Dichos Resultados Preliminares se realizaron con el fin de poder estimar la Pérdida Esperada (PE) por operación, la cual se calcula haciendo uso de la ecuación (1).

Los Resultados de la estimación de la PE se realizaron en el software estadístico R. El cuadro VI muestra los resultados obtenidos para las primeras 5 observaciones, los resultados para la base de datos con la que se esta trabajando están disponibles en el repositorio de Github. [7]

Identificación	EC	M	LGD	PD	PE
1	21437.57	1.000	100 %	0.00485	103.97
2	31296.89	1.050	100 %	0.01330	437.06
3	15967.01	1.000	100 %	0.02809	448.51
4	18110.20	1.025	100 %	0.07334	1361.41
5	4343.25	1.000	100 %	0.05480	238.01

Cuadro VI
RESULTADOS CENTRALES PÉRDIDA ESPERADA

IV-E2. Valor en Riesgo (VaR): Para el cálculo del VaR se utilizan las ecuaciones (2) y (8), donde la ecuación (8) es la Probabilidad Máxima de Incumplimiento la cual se calcula como una aproximación de la Distribución Normal Estándar $N(x, \mu = 0, \sigma = 1)$, los cálculos respectivos se realizaron en el software estadístico R; El cálculo de la probabilidad máxima de Incumplimiento requiere el uso de

las funciones $pnorm$ y $qnorm$ (Inversa de la distribución Normal Estándar). El cuadro VII muestra los resultados obtenidos de P_{max} para las primeras 5 operaciones, y el cuadro VIII muestra los resultados del VaR.

Los resultados para el total de operaciones con que cuenta la base de datos con la que se esta trabajando se encuentran disponibles en el repositorio de Github. [7]

PD	ρ	G(PD)	G(0.999)	Pmax
0.00485	0.1397	-2.5863	3.0902	0.06140
0.01330	0.1116	-2.2173	3.0902	0.10435
0.02809	0.0786	-1.9096	3.0902	0.13859
0.07334	0.0399	-1.4514	3.0902	0.19747
0.05480	0.0490	-1.6000	3.0902	0.17396

Cuadro VII
RESULTADOS CENTRALES PROBABILIDAD MÁXIMA DE INCUMPLIMIENTO

Desde el punto de vista de la gestión de Riesgo, el Valor en Riesgo es aquella pérdida máxima que se podría observar dado el nivel de confianza del 99,9 %.

Una interpretación de P_{max} para el cliente de la primera operación de crédito tiene una Probabilidad de Incumplimiento del 0.49 %, la cual, con el 99.9 % de confianza, podría variar desde 0 % hasta el 6.14 %.

De igual forma el quinto cliente tiene una Probabilidad de Incumplimiento del 5.48 %, la cual, con una confianza del 99,9 %, podría variar entre 0 % y 17.40 %

Identificación	EC	LGD	Pmax	VaR
1	21437.57	100 %	6.14 %	1316.29
2	31296.89	100 %	10.43 %	3265.96
3	15967.01	100 %	13.86 %	2213.01
4	18110.20	100 %	19.75 %	3576.26
5	4343.25	100 %	17.40 %	755.58

Cuadro VIII
RESULTADOS CENTRALES VALOR EN RIESGO

IV-E3. Pérdida Inesperada (PI): La Pérdida Inesperada o Capital en Riesgo se define en la ecuación (3) es la diferencia entre el Valor en Riesgo y la Pérdida Esperada, la estimación de la PI se realizó en el software estadístico R. En cuadro IX se muestran los resultados obtenidos para las primeras 5 operaciones, los Resultados obtenidos para el total de operaciones de la base de datos se encuentran disponibles en el repositorio de Github. [7]

Identificación	EC	PE	VaR	PI
1	21437.57	103.97	1316.29	1212.32
2	31296.89	437.06	3265.96	2828.90
3	15967.01	448.51	2213.01	1764.50
4	18110.20	1361.41	3576.26	2214.85
5	4343.25	238.01	755.58	517.57

Cuadro IX
RESULTADOS CENTRALES PÉRDIDA INESPERADA

La interpretación simultánea de los tres indicadores de Riesgo Financiero permite comprender mejor la relación entre

ellos.

La primera operación de crédito, dadas las condiciones actuales del cliente, podría representar una pérdida que se situaría entre Lps. 0 y Lps. 1,316.29, siendo la Pérdida Esperada Lps. 103.97.

La Institución Financiera debería realizar una estimación de crédito (provisión, previsión) equivalente a Lps. 103.97 y constituir capital de riesgo por Lps. 1,212.32.

El cliente de la operación 4 representa una pérdida que podría situarse entre cero y Lps. 3,576.26, siendo la Pérdida Esperada Lps. 1,361.41. El capital en riesgo llegaría a Lps. 2,214.85.

Pérdida Esperada	Valor en Riesgo	Pérdida Inesperada
3,534,354.3776	15,725,309.07	12,190,954.83

Cuadro X
INDICADORES DE RIESGO FINANCIERO (TOTAL DEL PORTAFOLIO)

IV-E4. Resultados Matriz de Transición (Base de Datos Año 2019): En los cuadros XI y XII se muestra el resumen obtenido al generar la Matriz de Transición para la base de datos correspondiente al Portafolio de Créditos del año 2019, la cual tiene un horizonte de tiempo de 12 meses y un horizonte de transición de 11 meses, el horizonte de tiempo inicia en el mes de Enero de 2019 y finaliza en el mes de Diciembre de 2019. La Matriz de Transición completa esta disponible en el repositorio de Github. [7]

Las Instituciones Financieras están interesadas en conocer el porcentaje de créditos que tuvieron una Recuperación al finalizar el año comercial, así mismo el porcentaje de créditos que tuvieron un Deterioro al transcurrir el año, es decir el porcentaje de créditos que empezaron en un estado de días de mora y al transcurrir el tiempo migraron a otro estado de días de mora mayor al que tenían inicialmente. En el cuadro XI comparamos el momento inicial (Días de Mora Enero) contra el momento final (Días de Mora Diciembre) del total del Portafolio de Créditos del año 2019.

Notamos que la Recuperación de Créditos a cero días de mora es del 86.9 % con un Deterioro crediticio del 13.09 %, el estado crediticio de un día de mora a treinta días de mora tuvo una Recuperación del 56.13 %, una Permanencia de 20.93 % correspondiente al total de créditos que empezaron en el estado un día de mora a treinta días de mora y al finalizar el horizonte de transición terminaron en un día de mora a treinta días de mora, el Deterioro crediticio es del 22.93 % correspondiente al total de créditos que empezaron en el estado un día de mora a treinta días de mora y al transcurrir el tiempo migraron a una calificación crediticia mayor a la que iniciaron. El mismo análisis se realiza para cada uno de los estados crediticios descritos en el cuadro XI.

Anteriormente se realizó el análisis para la Matriz de Transición comparando el momento inicial contra el momento final, pero algunas Instituciones Financieras les interesa conocer que realmente está pasando durante el horizonte de transición; Por lo que dicho análisis se puede realizar comparando momento inicial contra mora máxima, que corresponde a los días de mora mayor que tuvieron los créditos durante el horizonte de transición. Los Resultados de la Matriz de Transición resultante de comparar momento inicial contra mora máxima se encuentran resumidos en el cuadro XII.

Podemos observar en el cuadro XII que el porcentaje de créditos que tuvieron una Recuperación crediticia a cero días de mora es del 35.00 % y hubo un Deterioro Crediticio del 65.00 % que corresponde al total de créditos que empezaron en el estado crediticio cero días de mora y al transcurrir el tiempo migraron a un estado crediticio de días de mora mayor al que tenían inicialmente, el estado crediticio un día de mora a treinta días de mora tiene una Recuperación del 6.00 %, una Permanencia de 22.67 % y un Deterioro crediticio del 71.33 %.

Notamos que al comparar Momento Inicial contra Mora Máxima, el Deterioro crediticio aumenta de manera significativa que cuando se compara Momento Inicial contra Momento Final; Por lo que el análisis que cada Institución Financiera decida realizar va a depender de los resultados que desea obtener.

IV-E5. Resultados Matriz de Transición (Base de Datos Año 2020): En los cuadros XIII y XIV se muestra el resumen obtenido al generar la Matriz de Transición para la base de datos correspondiente al Portafolio de Créditos del año 2020, la cual tiene un horizonte de tiempo de 12 meses y un horizonte de transición de 11 meses, el horizonte de tiempo inicia en el mes de Febrero del 2020 y finaliza en el mes de Febrero del 2021. La Matriz de Transición completa esta disponible en el repositorio de Github. [7]

El análisis de los resultados obtenidos en los cuadros XIII y XIV es similar al análisis que se dio a los resultados obtenidos al generar la Matriz de Transición correspondiente a la base de datos del año 2019.

Notamos en el cuadro XIII (Momento Inicial - Momento Final) que la Recuperación crediticia a cero días de mora es del 76.30 % y el Deterioro crediticio es del 23.70 %, en el cuadro XIV (Momento Inicial - Mora Máxima) la Recuperación Crediticia a cero días de mora es del 21.10 % y el Deterioro crediticio es del 78.90 %, nuevamente es evidente que el Deterioro crediticio aumenta de manera significativa al comparar momento inicial contra mora máxima, esto se debe que al calcular la mora máxima se está tomando la mayor cantidad de días de mora durante el horizonte de transición para cada uno de los créditos del Portafolio.

Estado	Recuperación	Permanencia	Deterioro	Total	Estado	Recuperación	Permanencia	Deterioro	Total
000-000	86.91 %	0.00 %	13.09 %	100.00 %	000-000	21.10 %	0.00 %	78.90 %	100.00 %
001-030	56.13 %	20.93 %	22.93 %	100.00 %	001-030	3.55 %	3.55 %	92.91 %	100.00 %
031-060	57.81 %	11.98 %	30.21 %	100.00 %	031-060	1.45 %	8.70 %	89.86 %	100.00 %
061-090	57.69 %	10.58 %	31.73 %	100.00 %	061-090	3.70 %	11.11 %	85.19 %	100.00 %
091-120	56.14 %	15.79 %	28.07 %	100.00 %	091-120	0.00 %	0.00 %	100.00 %	100.00 %
121-150	61.11 %	11.11 %	27.78 %	100.00 %	121-150	0.00 %	12.50 %	87.50 %	100.00 %
151-180	62.50 %	8.33 %	29.17 %	100.00 %	151-180	16.67 %	0.00 %	83.33 %	100.00 %
181-210	64.29 %	7.14 %	28.57 %	100.00 %	181-210	0.00 %	0.00 %	100.00 %	100.00 %
211-240	57.89 %	0.00 %	42.11 %	100.00 %	211-240	0.00 %	0.00 %	100.00 %	100.00 %
241-270	75.00 %	12.50 %	12.50 %	100.00 %	241-270	0.00 %	0.00 %	100.00 %	100.00 %
271-300	66.67 %	0.00 %	33.33 %	100.00 %	271-300	0.00 %	0.00 %	100.00 %	100.00 %
301-330	50.00 %	0.00 %	50.00 %	100.00 %	301-330	0.00 %	0.00 %	100.00 %	100.00 %
331-360	25.00 %	0.00 %	75.00 %	100.00 %	Mayor a 360	0.00 %	0.00 %	100.00 %	100.00 %
Mayor a 360	35.90 %	64.10 %	0.00 %	100.00 %					

Cuadro XI

RESULTADOS CENTRALES MATRIZ DE TRANSICIÓN (MOMENTO INICIAL - MOMENTO FINAL (AÑO 2019))

Cuadro XIV

RESULTADOS CENTRALES MATRIZ DE TRANSICIÓN (MOMENTO INICIAL - MORA MÁXIMA (AÑO 2020))

Estado	Recuperación	Permanencia	Deterioro	Total
000-000	35.00 %	0.00 %	65.00 %	100.00 %
001-030	6.00 %	22.67 %	71.33 %	100.00 %
031-060	6.77 %	10.94 %	82.29 %	100.00 %
061-090	15.38 %	15.38 %	69.23 %	100.00 %
091-120	12.28 %	12.28 %	75.44 %	100.00 %
121-150	20.37 %	9.26 %	70.37 %	100.00 %
151-180	29.17 %	4.17 %	66.67 %	100.00 %
181-210	10.71 %	0.00 %	89.29 %	100.00 %
211-240	10.53 %	5.26 %	84.21 %	100.00 %
241-270	0.00 %	0.00 %	100.00 %	100.00 %
271-300	0.00 %	0.00 %	100.00 %	100.00 %
301-330	16.67 %	0.00 %	83.33 %	100.00 %
331-360	0.00 %	0.00 %	100.00 %	100.00 %
Mayor a 360	0.00 %	100.00 %	0.00 %	100.00 %

Cuadro XII

RESULTADOS CENTRALES MATRIZ DE TRANSICIÓN (MOMENTO INICIAL - MORA MÁXIMA (AÑO 2019))

Estado	Recuperación	Permanencia	Deterioro	Total
000-000	76.30 %	0.00 %	23.70 %	100.00 %
001-030	48.23 %	22.70 %	29.08 %	100.00 %
031-060	42.03 %	15.94 %	42.03 %	100.00 %
061-090	37.04 %	3.70 %	59.26 %	100.00 %
091-120	30.77 %	0.00 %	69.23 %	100.00 %
121-150	62.50 %	0.00 %	37.50 %	100.00 %
151-180	66.67 %	0.00 %	33.33 %	100.00 %
181-210	66.67 %	0.00 %	33.33 %	100.00 %
211-240	50.00 %	0.00 %	50.00 %	100.00 %
241-270	50.00 %	0.00 %	50.00 %	100.00 %
271-300	0.00 %	0.00 %	100.00 %	100.00 %
301-330	100.00 %	0.00 %	0.00 %	100.00 %
Mayor a 360	0.00 %	100.00 %	0.00 %	100.00 %

Cuadro XIII

RESULTADOS CENTRALES MATRIZ DE TRANSICIÓN (MOMENTO INICIAL - MOMENTO FINAL (AÑO 2020))

V. CONCLUSIONES

Las siguientes conclusiones están basadas en los resultados obtenidos al trabajar con las bases de datos (*baseRLB.xlsx*, *MT2019.xlsx*, *MT2020.xlsx*) disponibles en el repositorio de Github. [7]

1. Al aplicar el Test de Wald hacia atrás con fin de conocer el Contraste de Hipótesis de Nulidad de los Parámetros se puede cometer el Error de Tipo II, por lo que es importante el conocimiento previo que se tiene de las variables de tipo social que se introducen al realizar el Test de Wald para no descartar variables que realmente pueden ser explicativas en el modelo de Regresión Logística Binaria utilizado para el cálculo de la Probabilidad de Incumplimiento.
2. En el Modelo de Regresión Logística Binaria los Indicadores de Riesgo Financiero (PE, VaR, PI) se analizan para cada Cliente (Operación Crediticia) es decir desde un punto de vista Micro; Mientras que en el Modelo utilizando Cadenas de Markov (Matrices de Transición) el análisis se realiza a nivel de Cartera de Créditos (Portafolio) es decir a nivel Macro, por esta razón cada Institución Financiera dependiendo de sus intereses puede decidir cual modelo se ajusta mejor a los resultados que desea obtener.

VI. AUTOCRÍTICA

En general se han realizado una gran cantidad de metodologías para cuantificar las pérdidas a las que se expone cualquier empresa que tenga un portafolio de crédito o inversiones, dentro de estas metodologías, las que tienen mayor relevancia en la actualidad son aquellas que están fundamentadas bajo un modelo matemático estadístico. Sin embargo, el estudio de un modelo estadístico como ser las Matrices de Transición necesitan tener una data histórica para lograr que sus estimadores sean más precisos, muchos de estos estudios sugieren tener una data histórica de 12 meses, por lo menos para que los cálculos de la Probabilidad de Incumplimiento sean mas confiables, es importante señalar que dentro de este modelo solo se captura la transición de una operación bajo el supuesto de los días en mora, lo que nos obliga a conocer a profundidad el negocio o las condiciones en las que está el portafolio dado que algunos tienen un comportamiento estacional, tasa de intereses diferentes, diferentes tipos de garantías, etc.

Por otro lado, el modelo de Regresión logística Binaria es un modelo mas preciso que captura no solo el peso o la influencia de los días de mora si no que relaciona el cálculo de la Probabilidad de Incumplimiento con la información personal demográfica, laboral, económica, financiera, social, de cada uno de los clientes. Esto nos permite hacer estimaciones sin una data histórica y crear planes de negocios partiendo de las características del cliente y no de las operaciones. Sin embargo, la aplicación de esta metodología necesita tener bastante información sensible y actualizada del cliente o empresa, información que es muy difícil que los clientes o empresas proporcionen ya sea por seguridad o por incomodidad, también es imperativo mencionar que en este tipo de modelo el investigador debe de evitar muchos errores ya sean de muestreo, correlación, colinealidad y dependencia de las variables en estudio.

La diferencia entre el modelo de Matrices de Transición y Regresión Logística Binaria es que las Matrices de Transición capturan las Probabilidades de Incumplimiento por portafolio en una manera mejor dicha es un modelo Macro que a diferencia de la Regresión Logística Binaria que captura las Probabilidades de Incumplimiento por cliente o crédito lo convierte en modelo Micro. Como en todo modelo estadístico es importante que el investigador conozca a profundidad los portafolios con los que va a trabajar porque a partir de esa información y conocimiento sabrá que modelo se ajusta mejor a sus estimaciones y así lograr una mejor precisión en su estimación para que la empresa o negocio pueda crear planes de negocio, planes de reservas y poder mitigar de una manera confiable su riesgo de crédito.

VII. DISPONIBILIDAD DE LOS DATOS

Debido a la confiabilidad de la información personal de los usuarios financieros es casi imposible obtener datos reales que se puedan utilizar para realizar estudios en tesis académicas, no obstante, en algunos casos se pueden obtener bases de datos

que se ajusten a los parámetros que el investigador desear estimar. Para esta investigación los datos fueron adquiridos bajo unas series de talleres y cursos académicos que mis colegas de carrera han ido recibiendo con el pasar del tiempo por ende no se podrá decir que los datos utilizados para esta investigación reflejan la realidad de lo que esta sucediendo con el sistema financiero del país. Los datos mencionados anteriormente estarán disponibles en la plataforma Github que se podrá acceder con el siguiente link: <https://github.com/SM-Gomez/Seminario-de-Investigacion>. [7]

VIII. AGRADECIMIENTOS

En primer lugar quiero agradecer Dios y a mi supervisor Doctor: Fredy Antonio Vides Romero, quien con sus conocimientos y apoyo me guió a través de cada una de las etapas de este proyecto de investigación para alcanzar los resultados que buscaba.

También quiero agradecer a el Licenciado en Matemáticas con Orientación en Estadística: Kevin Kosner Pineda Castillo por brindarme todos los recursos y herramientas que fueron necesarios para llevar a cabo el proceso de investigación. No hubiese podido llegar a estos resultados de no haber sido por su incondicional ayuda.

Por ultimo quiero agradecer a mis padres y hermanos, por apoyarme en todo momento. En especial, a mis padres; Licenciada en Administración de Centros Educativos: María Emperatriz Meza, Perito Mercantil y Contador Publico: Quintin Gómez, quienes siempre han estado ahí para brindarme palabras de apoyo y un abrazo reconfortante.

REFERENCIAS

- [1] De La Fuente, M. (2004). Modelos de pérdida esperada. A. Elizondo, Medición integral del riesgo de crédito, 43-78.
- [2] Alonso Cifuentes, J. C., & Berggrun Preciado, L. (2008). Introducción al análisis de riesgo financiero.
- [3] Támara-Ayús, A., Aristizábal, R., & Velásquez, E. (2012). Matrices de transición en el análisis del riesgo crediticio como elemento fundamental en el cálculo de la pérdida esperada en una institución financiera colombiana. Revista Ingenierías Universidad de Medellín, 11(20), 105-114.
- [4] Ortega, M. F. METODOLOGÍA UTILIZADA PARA EL CÁLCULO DE RIESGO CRÉDITO EN UN PORTAFOLIO EFICIENTE DE CRÉDITO AL CONSUMO.
- [5] Cadenas de Markov
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/PEst/tema4pe.pdf>
- [6] Regresión Logística Simple y Múltiple.
https://rpubs.com/Joaquin_AR/229736
- [7] Github
<https://github.com/SM-Gomez>
- [8] Novalés, A. (2014). Valor en riesgo. Departamento de Economía Cuantitativa.

-
- [9] Comité de Supervisión Bancaria de Basilea (Basilea III: MARCO REGULADOR INTERNACIONAL PARA LOS BANCOS) https://www.bis.org/bcbs/basel3_es.htm
- [10] Berlanga Silvente, V., & Vilà Baños, R. (2014). Cómo obtener un modelo de regresión logística binaria con SPSS. REIRE. Revista d'Innovació i Recerca en Educació, 2014, vol. 7, num. 2.
- [11] Flores Manrique, Luz, Análisis Estadístico de los factores de riesgo que influyen en la enfermedad Angina de Pecho. https://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/flores_ml/Cap4.pdf
- [12] Modelo de Regresión Logística, <http://www.ugr.es/~mvargas/2.DiaNocioRegLogistica.pdf>
- [13] Cuevas Véliz, V. D. (2019). Desafíos de la industria bancaria en la estimación de Loss Given Default.