

# EL ANÁLISIS DE SUPERVIVENCIA EN LA GESTIÓN DEL RIESGO DE CRÉDITO: EL CASO DE UNA CARTERA HIPOTECARIA ECUATORIANA.

**Survival analysis in credit risk management: the case of an Ecuadorian mortgage portfolio.**

<sup>1</sup>Luis Felipe Guevara , <sup>1</sup>Miguel Flores , <sup>2</sup>Ana Cabezas-Martínez\*

<sup>1</sup>Escuela Politécnica Nacional, Facultad de Ciencias, Departamento de Matemáticas, MODES / SIGTI, Quito, Ecuador.

<sup>2</sup>Facultad Latinoamericana de Ciencias Sociales, Departamento de Estudios Políticos, Quito, Ecuador.

\*ana.cabezas90@gmail.com

## Resumen

El presente documento pretende analizar la aplicación del análisis de supervivencia, y específicamente del modelo de cox en un ámbito como el financiero, mismo que no ha sido tradicionalmente abordado en América Latina. Para ello, en primera instancia se ha descrito y analizado la bibliografía existente con respecto a los beneficios que la aplicación del análisis de supervivencia deriva en la dinámica del sistema financiero formal, posteriormente se ha propuesto un procedimiento metodológico que permita aplicar un análisis de supervivencia a una cartera de créditos hipotecario pertenecientes a una entidad financiera ecuatoriana, con el propósito de segmentarla e identificar situaciones internas y externas de cara a la definición de estrategias diferenciadas tanto para la concesión crediticia como para la gestión de cobranzas. Finalmente, se pretendió automatizar el análisis realizado a través de la programación en el software R y con ello facilitar la reproducción y parametrización del mismo de cara a su implementación como parte de las acciones de gestión y mitigación del riesgo de crédito asociado a la cartera hipotecaria analizada.

**Palabras claves:** Dinámica del sistema financiero, concesión crediticia, análisis de riesgo.

## Abstract

This document aims to analyze the application of survival analysis, and specifically the cox model in an area such as finance, which has not traditionally been addressed in Latin America. In the first instance the existing bibliography has been described and analyzed with respect to the benefits that the application of the analysis of survival derives in the dynamics of the formal financial system, later a methodological procedure has been proposed that allows to to apply a survival analysis to a mortgage portfolio loans belonging to an Ecuadorian financial institution, with the purpose of segmenting it and identifying internal and external situations in order to define differentiated strategies for both the credit granting and the collection management. Finally, the analysis was automated through programming in the R software, thus facilitating its reproduction and parameterization for its implementation as part of the credit risk management and mitigation actions associated with the mortgage portfolio analyzed.

**Keywords:** Financial system dynamics, credit granting, risk analysis.

**Fecha de recepción:** 15-05-2021

**Fecha de aceptación:** 03-09-2021

**Fecha de publicación:** 13-10-2021

## I. INTRODUCCIÓN

El sistema financiero, en tanto se constituye en un esquema organizado que traslada liquidez desde unidades excedentarias hacia unidades de-

ficitarias, aporta significativamente en la transformación del ahorro en inversión, y con ello al crecimiento económico de un país, dependiendo de las particularidades de cada una de las economías. (1)

En esta línea, la sostenibilidad de las instituciones financieras, es un elemento relevante en el buen funcionamiento del engranaje denominado sistema financiero. Una de las aristas que contribuyen a este fin, es el adecuado manejo del riesgo de crédito, entendido como la administración de la posibilidad de pérdida suscitada por el no pago de obligaciones acordadas (2).

Entre los procesos planteados para la administración y gestión del mencionado riesgo, se realizan actividades de revisión permanente del portafolio de créditos, las cuales buscan pre-identificar situaciones de mayor probabilidad de incumplimiento o deterioro de cartera a través del monitoreo de diferentes variables (3).

Para ello, a través de diferentes técnicas estadísticas las diferentes entidades buscan estimar la probabilidad de que un cliente pueda ser buen o mal pagador, de acuerdo a sus particularidades, en el mejor de los casos, a través de la implementación de mecanismos de calificación estadística (scoring) que utilizan modelos binarios logit o probit, y así mejorar los procesos de selección de clientes y posteriormente de gestión de cobranzas (4).

Por otra parte, se puede mencionar que, con el devenir del tiempo, el control de calidad, y específicamente el control estadístico de la calidad, se ha constituido en una práctica, inicialmente aplicada en el sector industrial, que se ha extendido hacia los diferentes sectores productores de bienes y servicios, públicos y privados (5).

El enfoque de control estadístico de calidad para monitorear el desempeño de un proceso cobra relevancia, ya que a través de la sistematización del mismo se puede tener una visión global y simultánea de los resultados del proceso, haciendo factible la retroalimentación del mencionado de forma oportuna (6).

En esta lógica se inserta el denominado “análisis de supervivencia”, cuyos inicios se produjeron a través de la elaboración de las tablas de mortalidad atribuidas al astrónomo Edmon Halley quien utilizó en primera instancia un registro de funerales y nacimientos en la ciudad de Breslau. Tal como se lo conoce hoy, el análisis de supervivencia se lo aplicó inicialmente en la ingeniería, encaminado a analizar la duración y fiabilidad de los elementos que forman una máquina (7).

Puntualmente, el análisis de supervivencia es un

conjunto de procedimientos estadísticos que posibilitan el estudio de los datos y cuya variable de interés primordial es el tiempo de duración de un evento (8).

La problemática que pretende abordar la presente investigación aplicada, es develar los beneficios que la introducción del análisis de supervivencia, derivado del enfoque de control estadístico de calidad (8), puede detonar en el análisis y administración del riesgo de crédito en entidades financieras, y puntualmente en el caso de la cartera hipotecaria generada por una entidad financiera ecuatoriana.

Específicamente, mediante las técnicas estadísticas tradicionales se estima la probabilidad de pago y no pago de un cliente en una entidad financiera, pero no se identifica y monitorea el tiempo que transcurre desde la concesión del crédito hasta un posible evento de no pago y las razones que motivan ese comportamiento (9).

Es así, que a través de un ejercicio empírico, utilizando el análisis de supervivencia y particularmente el modelo de Cox (semi-paramétrico) se pretende destacar los beneficios que se suscitan de incorporar al tratamiento tradicional del riesgo de crédito el análisis de la temporalidad de no pago, ya que con ello se puede implementar con mayor certeza y de forma anticipada políticas de concesión y gestión crediticia que permitan mantener saludable la cartera de créditos de las diferentes entidades financieras (9). Cabe resaltar, que, en la bibliografía existente, la aplicación de este tipo de modelos en el sector financiero aún es incipiente a nivel de América Latina y obviamente en el Ecuador, por lo cual se torna de interés la realización de investigación aplicada en este campo (10).

Pese a lo mencionado, existe bibliografía de planteamientos y aplicación del análisis de supervivencia vinculado al Sistema Financiero, principalmente, por fuera de América Latina, lo cual denota la factibilidad real de implementar este tipo de análisis a nivel Ecuador. Ejemplo de lo expuesto ha sido la aplicación del análisis de supervivencia con información de clientes de crédito de consumo en un banco Español (11), con información de créditos hipotecarios pertenecientes a minoristas en Estados Unidos (12), con información de créditos personales derivados de las instituciones que ofertan servicios financieros

en Australia (13), con información de créditos generados en el Rabobank Internacional situado en Holanda (14) y con información de créditos corporativos y grandes empresas en Perú entre algunos destacados (10).

Resumiendo, algunos elementos relevantes derivados de las experiencias de análisis y aplicación del análisis de supervivencia en el ámbito del sistema financiero se pueden destacar los siguientes:

- La principal ventaja del análisis de supervivencia es la capacidad de incorporar datos censurados a la información sujeta de análisis (14).
- Existen ejercicios de comparación de técnicas que denotaron que la transformación de logrank (modelo de supervivencia de cox) supera a la transformación logística y otros, en cuanto a la predicción de la probabilidad de supervivencia, principalmente basados en la prueba estadística de Wald (14).
- La identificación de la temporalidad en la cual se puede incurrir en incumplimiento (default) es un elemento importante de cara a mitigar el riesgo de crédito, para lo cual el típico análisis de matrices de transición es insuficiente. Así también se vislumbró la necesidad de incorporar a este tipo de modelos variables asociadas al entorno económico y variables microeconómicas específicas de los deudores (10-15).
- La censura da lugar a errores en las estimaciones de los parámetros para todas las metodologías (16). Es por ello, que en algunos casos se tomaron sub-bases de datos de la base general (créditos que no cumplen con una maduración mínima) con la finalidad de evitar casos censurados innecesarios (9).

Finalmente, en virtud de lo expuesto previamente, la hipótesis a testear en el presente trabajo de investigación aplicada es la siguiente:

“La aplicación e implementación de un modelo de supervivencia conllevaría beneficios de cara a la gestión del riesgo de crédito de una cartera hipotecaria ecuatoriana a través del análisis de la temporalidad en la que podrían suscitarse los “no pagos”, de la identificación de las variables que incidirían en el comportamiento referido; y de la implementación de políticas de crédito y gestión que permitan mitigar anticipadamente los riesgos identificados”

## II. MATERIALES Y MÉTODOS

Metodológicamente, en primera instancia, se elaborará un análisis de supervivencia, mediante el modelo no paramétrico de Kaplan y Meier y particularmente mediante el modelo de Cox para una base de créditos hipotecarios perteneciente a una entidad financiera ecuatoriana. El evento de interés que se constituirá en la variable de interés será la mora mayor a 90 días. Es así que la función de supervivencia a estimar estará relacionada con el tiempo que existe desde el desembolso crediticio hasta la ocurrencia de una mora grave. El procedimiento metodológico que se aplicará será el siguiente:

- Definición de la variable de interés.
- Breve descripción de la base de datos a utilizar.
- Análisis descriptivo de las variables consideradas inicialmente para formar parte del modelo de cox y construcción individual de funciones de supervivencia mediante el método de Kaplan y Meier.
- Formulación del modelo de Cox.
- Verificación de los supuestos del modelo de Cox.
- Interpretación de los resultados
- Segmentación de la base de clientes de la cartera de créditos hipotecarios, a partir del análisis de supervivencia desarrollado con la finalidad de plantear estrategias diferenciadas tanto de concesión como de gestión de cobranzas.

Adicionalmente, se automatizará el análisis de supervivencia a través de la programación del análisis en el software R, con lo cual se podrá facilitar la reproducción y parametrización del análisis de cara a la implementación del análisis como parte de las acciones dentro de la mitigación y gestión de los riesgos de crédito asociados a la cartera hipotecaria analizada.

En lo que respecta al marco teórico del presente trabajo a continuación se detallan el estimador de Kaplan y Meier y el modelo de Cox que forman parte de las técnicas utilizadas dentro del análisis de supervivencia.

### Estimador de Kaplan y Meier

El estimador Kaplan-Meier desarrollado en 1958 es un método no paramétrico que estima la probabilidad de supervivencia  $\hat{S}(t)$  en el momento en el que se ha suscitado el evento. Para ello, es im-

portante tener en cuenta que para sobrevivir en un momento determinado  $t$ , se tiene que haber sobrevivido en todos los instantes anteriores a  $t$  (8).

El cálculo de este método parte de una probabilidad condicional compuesta, en la cual se multiplica la supervivencia en el instante anterior y la tasa de supervivencia en ese instante. El detalle explícito del cálculo de la probabilidad de supervivencia bajo este método sería el siguiente (17):

$$\hat{S}(t) = \prod_{j: \tau_j \leq t} \left[ 1 - \frac{d_j}{n_j} \right] \quad (1)$$

Donde  $d_j$  es el número de individuos a los cuales les suscitó el evento de interés en el tiempo  $t_j$ ,  $n_j$  es el número de individuos que se encontraban en riesgo de que les suscitara el evento de interés. Adicionalmente  $n_j$  es la muestra original de individuos menos aquellos que fueron censurados o aquellos a los que les suscitó el evento de interés antes del tiempo  $t_j$  (18).

### Modelo de Cox

**Definición y Funciones características:** Es una forma funcional utilizada en el análisis de supervivencia conocido también como modelo de riesgo proporcional (desarrollada en 1972), que hace uso de una propuesta semi-paramétrica para encontrar una función que describa el riesgo de ocurrencia de un evento en diferentes períodos de tiempo (19).

El modelo se asemeja a una regresión que describe la relación entre la ocurrencia de un evento de interés o falla y un conjunto de variables. Formalmente, cuando  $t \geq 0$  la referida función de riesgo luce de la siguiente manera (9-20).

Donde  $B'x$  es el vector de covariables, la parte paramétrica  $e^{(B'x)}$  se la conoce como “puntaje de riesgo” y la parte no paramétrica  $h_0(t)$  se la conoce como “función de riesgo base” o aquella que describe la probabilidad de ocurrencia del evento cuando todas las variables explicativas son 0.

Algunos de las características relevantes del modelo de Cox son (21). Es producto de una función en  $t$  y una función en  $x$ ;

Algunos de las características relevantes del modelo de Cox son (21). Es producto de una función en  $t$  y una función en  $x$ ;

- $x$  es independiente del tiempo;
- La función de riesgo base  $h_0(t)$  no depende de  $x$  sino exclusivamente del tiempo;
- El exponencial ( $e$ ) considera las  $x$  pero no  $t$ ;

- Sigue los supuestos del modelo de riesgo proporcional;

- Las estimaciones de riesgo son no negativas.

La función de supervivencia relacionada al modelo de Cox es la siguiente:

$$S(t|x) = e^{-H_0(t)e^{(B'x)}} \quad (2)$$

La función de distribución del modelo de Cox es la siguiente:

$$F(t|x) = 1 - e^{-H_0(t)e^{(B'x)}} \quad (3)$$

**Supuestos del Modelo de Riesgo Proporcional:** El supuesto clave del modelo proporcional es que el riesgo de un evento en un grupo es un múltiplo constante del riesgo en cualquier otro grupo (9). Así también se puede mencionar que las tasas de riesgo son proporcionales cuando la relación entre el riesgo de dos individuos con variables explicativas  $Z$  y  $Z^*$  es constante en el tiempo, siendo el riesgo calculado para cada predictor una proporción del riesgo base  $h_0$ . Este supuesto se verifica a través del comportamiento aleatorio de los residuos de Schoenfeld y de forma relacionada la prueba rho la cual sigue una distribución ji-cuadrada y cuya hipótesis nula es el cumplimiento del supuesto de riesgo proporcional reseñado (22).

Otros supuestos del modelo de riesgo proporcional son los siguientes (23-24):

- No existe influencia de los individuos en la estimación del modelo. Este supuesto se verifica mediante la identificación del comportamiento de los residuos denominados “*deviance*”, los cuales son una transformación de normalización de los residuos de martingala. La no identificación de casos atípicos en la distribución de estos residuos transformados refiere la no influencia de individuos en el modelo de supervivencia desarrollado. El cálculo de este tipo de residuos obedece a la siguiente fórmula:

$$d_i = \text{sgn}(\tilde{M}_i) * \sqrt{-\tilde{M}_i - N_i \log((N_i - \tilde{M}_i)/N_i)} \quad (4)$$

- No existe influencia de los individuos en la estimación de cada parámetro del modelo. Este supuesto se verifica mediante la identificación del comportamiento de los residuos denominados “*score*”. La no identificación de casos atípicos en la distribución de estos residuos refiere la no influencia de los coeficientes en el modelo de supervivencia desarrollado. Los residuos “*score*” se definen de la siguiente manera:

$$U_{ij} = U_{ij}(\hat{\beta}, \infty) \quad (5)$$

Donde  $U_{ij}(\beta, t), j=1, \dots, p$  son las componentes del



vector fila de longitud  $p$ , obtenido mediante el proceso de score para el  $i$ -ésimo individuo:

$$U_i(\beta) = \int_0^t [Z_i(s) - \bar{Z}(\beta, s)] dN_i(s) \quad (6)$$

• Se supone una adecuada forma funcional de cada variable continua. Este supuesto se verifica mediante la identificación del comportamiento de los residuos de martingala. El contraste de los residuos de martingala del modelo frente a los valores de la variable continua, permitirá vislumbrar si la asunción de linealidad se cumple. Los residuos de martingala se definen de la siguiente manera:

$$\hat{M}_i(t) = N_i(t) - \hat{E}_i(t) = N_i(t) - \int_0^t Y_i(s) e^{\beta' Z_i(s)} d\hat{R}_0(\beta, s) \quad (7)$$

Donde  $\hat{M}_i(t)$  se deriva de la martingala de un proceso de conteo para el  $i$ -ésimo individuo y se define como la diferencia entre el proceso de conteo  $N_i(t)$  y la integral de la función de intensidad (25). Dentro de la referida integral  $\hat{R}_0(\beta, s)$  es el estimador de riesgo de Breslow o de Nelson y Aalen.

### III. RESULTADOS

#### A. Desarrollo del Análisis de Supervivencia

*Definición de la Variable de Interés:* La variable de interés es el tiempo en meses que tarda un crédito hipotecario en caer en una mora grave, denominada “tiempoM90pv\_2mes” (tiempo de supervivencia). Se define como mora grave al hecho de permanecer en mora por más de 90 días consecutivos luego de la fecha en la cual debía hacerse el pago correspondiente. Explícitamente, la referida variable se condensa en una de tipo binaria denominada “Mayor90”, misma que toma el valor “1” si la operación de crédito hipotecario ha experimentado una mora grave y “0” si no lo ha experimentado durante el período de estudio (caso censurado por la derecha).

El vector conformado por las dos variables mencionadas en cada uno de los casos de los clientes del crédito hipotecario, se constituye en el tipo de variable dependiente requerido para la formulación de un análisis de supervivencia.

Datos: Los datos utilizados para la estimación fueron los créditos hipotecarios activos al 31 de diciembre de 2016 que denotaron una antigüedad de al menos 2 años y fueron destinados a la adquisición de vivienda terminada. Es decir que fueron concedidos entre los años 2008 y 2014. En virtud de lo anterior, el número de observacio-

nes consideradas fue de 79.225 observaciones.

Complementariamente, el tiempo de estudio definido para el presente análisis fue el comprendido entre el 30 de julio del 2008 y el 31 de diciembre de 2016, es decir un período de 102,53 meses u 8,54 años.

#### Caracterización de Variables – Modelo Kaplan – Meier:

##### a) Variable: Subregión Geográfica

La variable subregión geográfica (SubReg) es una variable nominal que agrupa en subregiones geográficas las provincias en las cuales los clientes obtuvieron su crédito hipotecario en seis (6) opciones de respuesta: “SIERRA\_N (CAR., IMB., PIC.)”, “SIERRA\_C (COT., TUNG., BOL., CHI)”, “SIERRA\_S (CAÑ., AZU., LOJ.)”, “COSTA (incluye STO.DOM. y GAL.)”, “ORIENTE\_N (SUC., NAP., ORE.)” y “ORIENTE\_CS (PAS, MOR., ZAM.)”.

SubReg	Mayor90		Row Total
	0	1	
SIERRA_N (CAR., IMB., PIC.)	28138	2442	30580
	92%	8%	38,60%
	39,40%	31,50%	
ERRA_C (COT., TUNG., BOL., CHI)	3368	214	3582
	94%	6%	4,50%
	4,70%	2,80%	
SIERRA_S (CAÑ., AZU., LOJ.)	4816	265	5081
	94,80%	5,20%	6,40%
	6,70%	3,40%	
COSTA (incluye STO. DOM. Y GAL)	33370	4483	37853
	88,20%	11,80%	47,80%
	47%	58%	
ORIENTE_N (SUC., NAP., ORE)	1169	298,00%	1467
	79,70%	20,30%	1,90%
	1,60%	3,80%	
ORIENTE_CS (PAS, MOR., ZAM.)	610	52	662
	92,10%	7,90%	0,80%
	0,90%	0,70%	
Column Total	71471	7754	79225
	90,20%	9,80%	

**Tabla 1.** Tabla de contingencia - variable subregión geográfica (SubReg) vs variable presencia mora grave (Mayor90)

Total Observations in Table: 79225

Statistics for All Table Factors

Pearson's Chi-squared test

Chi<sup>2</sup> 2+659.7878. d.f.=5. p=2.425538e-140

Minimum expected frequency: 64.79202

El análisis de frecuencia de la variable analizada

refirió en primera instancia no contar con casos perdidos, con lo cual contó con 79.225 casos. Así también develo que los clientes del crédito hipotecario considerados se distribuyeron según la variable subregión geográfica (SubReg) mayoritariamente en un 47,8% en la opción “COSTA (incluye STO.DOM. y GAL.)” y en un 38,6% en la opción “SIERRA\_N (CAR., IMB., PIC.)”.

A partir del análisis de la variable mora grave (Mayor90) según las opciones de la variable subregión geográfica (SubReg) se puede señalar que la proporción de quienes incurrieron en una mora grave se suscitó en mayor medida en las opciones “ORIENTE\_N (SUC., NAP., ORE.)” (20,3%) y “COSTA (incluye STO.DOM. y GAL.)” (11,8%), incluso superando levemente la estructura natural de la opción presencia de mora grave (1) de la variable mora grave (Mayor90) (9.8%). Adicionalmente, a partir la prueba ji-cuadrado de independencia realizada a la tabla de contingencia de las variables subregión geográfica (SubReg) y mora grave (Mayor90) se puede señalar que dado un p-valor asociado inferior a 0,05 ( $2,425538e-140$ ) se puede rechazar la hipótesis de independencia de variables ( $H_0$ ) y concluir que existe una relación de dependencia entre las dos variables analizadas.

Por otra parte, la aplicación del análisis de supervivencia mediante el modelo de Kaplan – Meier (KM) para identificar la función de supervivencia frente a la posibilidad de incurrir en una mora grave (Mayor 90) según la variable subregión geográfica (SubReg), denotó que la probabilidad de supervivencia tendió a descender en mayor medida a partir del mes 20, y con mayor intensidad en aquellos clientes localizados en las subregiones “ORIENTE\_N (SUC., NAP., ORE.)” y “COSTA (incluye STO.DOM. y GAL.)”

Así también, cabe indicar que dada la alta proporción de casos censurados en cada una de las subregiones geográficas no fue posible calcular la mediana de supervivencia en cada uno de los casos.

	n	events	median
SubReg=SIERRA_N (CAR., IMB., PIC.)	30580	2442	NA
SubReg=SIERRA_C (COT., TUNG., BOL., CHI)	3582	214	NA
SubReg=SIERRA_S (CAÑ., AZU., LOJ)	5081	265	NA

SubReg=COSTA (incluye STO. DOM. Y GAL.)	37853	4483	NA
SubReg=ORIENTE_N (SUC., NAP., ORE.)	1467	298	NA
SubReg=ORIENTE_CS (PAS,MOR., ZAM)	662	52	NA
	0,95LCL	0,95UCL	NA
SubReg=SIERRA_N (CAR., IMB., PIC.)	NA	NA	
SubReg=SIERRA_C (COT., TUNG., BOL., CHI)	NA	NA	
SubReg=SIERRA_S (CAÑ., AZU., LOJ)	NA	NA	
SubReg=COSTA (incluye STO. DOM. Y GAL.)	NA	NA	
SubReg=ORIENTE_N (SUC., NAP., ORE.)	NA	NA	
SubReg=ORIENTE_CS (PAS,MOR., ZAM)	NA	NA	

Tabla 2. Tabla de función de supervivencia, método KM, variable subregión geográfica (SubReg).

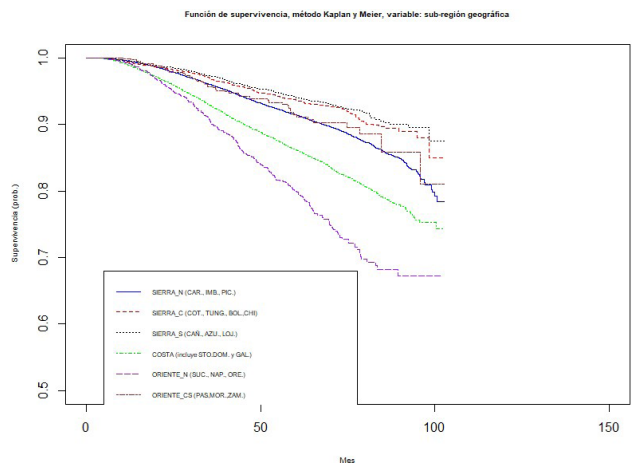


Figura 1. Función de supervivencia, método KM, variable subregión geográfica (SubReg).

Por su parte, luego de aplicar el estadístico de contraste log-rank el cual sigue una distribución ji-cuadrado para la evaluación de la hipótesis de igualdad de funciones de supervivencia a un nivel de confianzas del 95%, se puede afirmar que dado un p-valor cercano a 0, se rechaza la hipótesis nula ( $H_0$ ) de igualdad de funciones, y con ello se puede indicar que las funciones de supervivencia entre quienes integran las diferentes subregiones geográficas fueron diferentes.

Es importante señalar, que con la finalidad de utilizar esta variable en la formulación de un modelo multivariado se decidió agrupar sus opciones de respuesta en dos (2): “COSTA (incl. STO y GAL.) y ORIENTEN (NAP,SUC,ORE)” y “OTRAS SUBREGIONES”. El nuevo nombre de la variable con el ajuste citado sería SubReg2.

	N	Observed
SubReg=SIERRA_N (CAR., IMB., PIC.)	30580	2442
SubReg=SIERRA_C (COT., TUNG., BOL., CHI)	3582	214
SubReg=SIERRA_S (CAÑ., AZU., LOJ)	5081	265
SubReg=COSTA (incluye STO. DOM. Y GAL.)	37853	4483
SubReg=ORIENTE_N (SUC., NAP., ORE.)	1467	298
SubReg=ORIENTE_CS (PAS,MOR., ZAM)	662	52
	Expected	(O-E) √2/E
SubReg=SIERRA_N (CAR., IMB., PIC.)	3136	153.55
SubReg=SIERRA_C (COT., TUNG., BOL., CHI)	378	71.49
SubReg=SIERRA_S (CAÑ., AZU., LOJ)	519	124.37
SubReg=COSTA (incluye STO. DOM. Y GAL.)	3496	278.93
SubReg=ORIENTE_N (SUC., NAP., ORE.)	154	134.96
SubReg=ORIENTE_CS (PAS,MOR., ZAM)	71	5.11
	(O-E) <sup>2</sup> /V	
SubReg=SIERRA_N (CAR., IMB., PIC.)	258.44	
SubReg=SIERRA_C (COT., TUNG., BOL., CHI)	75.19	
SubReg=SIERRA_S (CAÑ., AZU., LOJ)	133.33	
SubReg=COSTA (incluye STO. DOM. Y GAL.)	509.97	
SubReg=ORIENTE_N (SUC., NAP., ORE.)	137.73	
SubReg=ORIENTE_CS (PAS,MOR., ZAM)	5.16	

**Tabla 3.** Resultado del contraste de igualdad de funciones de supervivencia para la variable subregión geográfica (SubReg).  
Chisq=771 on 5 degrees of freedom, p=0

La metodología desarrollada para el análisis de la variable “sub región geográfica”, se la realizó para las variables: rango de avalúo, género, rango de plazo, rango de ingreso, rango de monto financiado, rango de edad, actividad económica del empleador, ciclo económico al momento del desembolso y cesante previo a la mora grave.

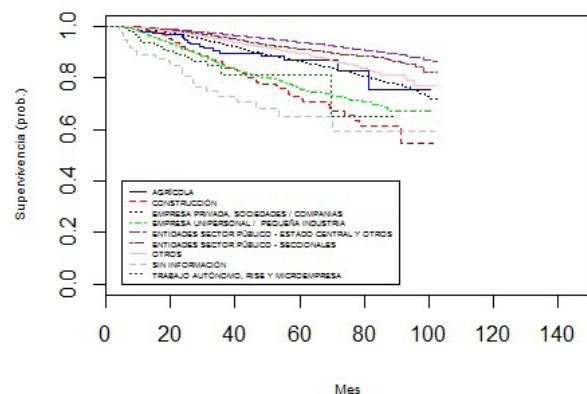
## B. Formulación de un Modelo de Regresión de Cox

1) *Consideraciones Específicas y Base de datos:* Conforme los análisis previos realizados con respecto a las variables a utilizar y sus correspondientes funciones de supervivencia construidas, se procedió a probar algunos modelos preliminares de COX utilizando la totalidad de la base de datos perfilada de manera previa (79.225 casos). En estos ejercicios se encontraron variables estadísticamente significativas, pero que aportaban poco en la explicación de la varianza de la tota-

lidad de los datos, en gran medida dado que las observaciones con mora mayor a 90 días apenas representaban el 9,8% de la totalidad de las observaciones.

En virtud de lo anterior, se decidió tomar una sub-base de datos para la formulación de un modelo de Cox más robusto. Específicamente, para la construcción del modelo se utilizó una sub-base de datos, construida con los casos de los clientes de créditos hipotecarios que desempeñaron actividades en los sectores “EMPRESA UNIPERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN”, dado que en los análisis previos y específicamente en las curvas de supervivencia construidas de la variable “actividad económica del empleador” se vislumbró que quienes se encontraban asociados a las actividades mencionadas experimentaban, en términos relativos, el descenso más pronunciado en el tiempo de su probabilidad de sobrevivir al evento de una mora grave.

Función de supervivencia, método Kaplan y Meier, variable: actividad económica emp.



**Figura 2.** Función de supervivencia, método KM, variable actividad económica (Acteco\_r)

Con lo expuesto anteriormente, la base utilizada para la construcción del modelo de COX fue aquella constituida por 3.570 casos que cumplen con las consideraciones expuestas en el párrafo anterior.

2) *Variables Utilizadas:* Como resultado de probar múltiples combinaciones de variables y de probar su validez y aporte, se definieron las siguientes como las idóneas entre las disponibles para integrar el modelo de cox propuesto.

### a) Variable Dependiente

- Variable continua tiempo en meses que tarda un crédito hipotecario en caer en una mora grave (tiempoM90pv\_2mes).

- Variable binaria mora grave (Mayor90). Opciones de respuesta: "0".- No ha experimentado mora grave y "1".- Experimentó una mora grave. Las variables expuestas configuraron la tipología de variables dependientes requeridas para estimar un modelo de supervivencia.

#### b) Variables Explicativas

En primera instancia, de las variables analizadas previamente se seleccionaron las siguientes:

- Variable nominal rango de ingreso (R\_Ingp). Opciones de respuesta: "hasta 770 USD (Q1)" (opción base), "entre 770,01 y 1.133 USD (Q2)", "entre 1.133,01 y 1.603 USD (Q3)", "entre 1.603,01 y 2.362 USD (Q4)" y "mayor a 2.362 USD (Q5)".
- Variable nominal rango de plazo (R\_Plazo). Opciones de respuesta: "menor a 20 años" (opción base) y "mayor a 20 años".
- Variable nominal subregión geográfica resumida (SubReg2). Opciones de respuesta: "otras subregiones" (opción base) y "Costa (incluye Sto. Dom. y Gal.) y Oriente Norte (Nap., Suc., Ore)".
- Variable nominal ciclo económico al momento del desembolso resumido (Ciclo2). Opciones de respuesta: "0".- desempeño económico sobre el crecimiento promedio de largo plazo (opción base) y "1".- desempeño económico bajo crecimiento promedio de largo plazo.

Adicionalmente, con la finalidad de mejorar la capacidad del modelo para explicar la variabilidad de los datos se incorporaron las siguientes variables:

- Variable discreta año de concesión (tyear).
- Variable nominal ciclo económico previo a la mora grave (Cicloef2). Opciones de respuesta: "0".- desempeño económico sobre el crecimiento promedio de largo plazo (opción base) y "1".- desempeño económico bajo crecimiento promedio de largo plazo.
- Variable nominal cesante y ciclo económico por debajo del crecimiento de largo plazo previo a la mora grave (cesantexciclof). Opciones de respuesta: "0".- El cliente no se encontró simultáneamente en la condición de cesante y en un ciclo económico por debajo del crecimiento de largo plazo previo a la mora grave (opción base) y "1".- El cliente sí se encontró simultáneamente en la condición de cesante y en un ciclo económico por debajo del crecimiento de largo plazo previo a la mora grave.

n=3570, number of events=695	
	coef
R_Ingpentre 770,01 y 1.133 USD (Q2)	0,377662
R_Ingpentre 1.133,01 y 1.603 USD (Q3)	0,384910
R_Ingpentre 1.603,01 y 2.362 USD (Q4)	0,535503
R_Ingpmayor a 2.362 USD (Q5)	0,891376
R_Plazomayor a 20 años	0,252111
SubReg2COSTA (incl. STO y GAL.) y ORIENTEN (NAP, SUC,ORE)	0,544825
Ciclo21	0,836868
tyear	1,178660
Cicloef21	-5,699248
cesantexciclof	2,039368

**Tabla 4.** Modelo de cox – Clientes de créditos hipotecarios que desempeñaron labores en "Empresa unipersonal / construcción / trabajo autónomo, rise y microempresa / sin información".

### 3) Evaluación e Interpretación de Variables del Modelo:

a) *Evaluación del Modelo:* Es importante señalar que en función de que el p-valor asociado al estadístico z en cada una de las variables explicativas fue inferior a 0,05, se rechazó la hipótesis nula de que cada uno de los coeficientes es igual a 0 ( $H_0$ ) y con ello se puede afirmar que todas las variables de manera individual resultaron estadísticamente significativas.

Por otra parte, los p-valor asociados a los estadísticos test de razón de verosimilitud (likelihood ratio test), test de wald (walt test) y test de puntajes (score [logrank]) al presentar magnitudes inferiores a 0,05, se puede rechazar la hipótesis de que los coeficientes del modelo en su conjunto son iguales a 0 ( $H_0$ ), y con ello se puede afirmar que el modelo en su conjunto es estadísticamente significativo.

Complementariamente, el coeficiente de determinación r-cuadrado del modelo obtenido indicó que el modelo puede explicar el 0,488 de la varianza de la problemática analizada. Si bien la magnitud del coeficiente pudo lucir insuficiente ante la necesidad de estimar los tiempos de supervivencia a partir de las variables contempladas en el modelo, cabe señalar que esta situación puede mejorar substancialmente con la incorporación de dos variables que no estuvieron disponibles al momento de la construcción del presente modelo como son: 1) peor calificación crediticia en los 2 años anteriores a la concesión del crédito y 2) estado civil del cliente, las cuales han sido relevantes en la identificación de casos de mora grave en ejercicios empíricos realizados.



Finalmente, en cuanto a la validación de los supuestos de riesgo proporcional se puede mencionar que los p-valores asociados al estadístico Rho, que sigue una distribución ji-cuadrado, pertenecientes a casi la totalidad de las variables y al modelo de forma global presentaron una magnitud superior 0,05, con lo cual no existiría evidencia significativa al 95% de confianza de que exista una violación al supuesto de riesgo proporcional, en cuanto a que las tasas de riesgo se mantengan constantes en el tiempo. Adicionalmente, la distribución relativamente aleatoria vislumbrada en los residuos “desviance” y “score” del modelo aplicado, evidencian la no existencia de influencia de los individuos en la estimación del modelo y de cada uno de los parámetros o coeficientes asociados. No se evaluó el supuesto adecuada forma funcional para variables continuas, ya que no se utilizaron variables de este tipo. Consecuentemente, de manera general se cumplen con los diferentes supuestos de riesgo proporcional.

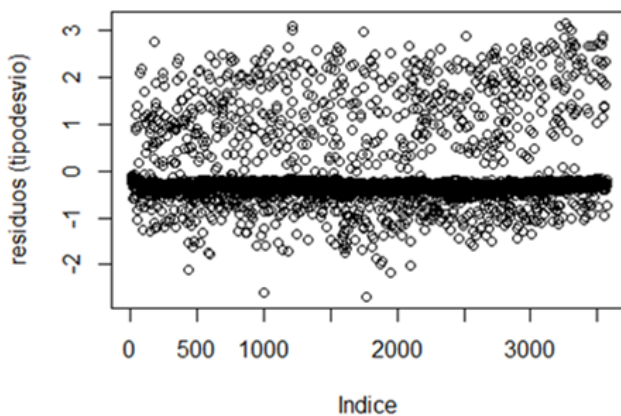


Figura 3. Residuos tipo desviance del modelo

b) *Interpretación de Resultados:* En primera instancia, a partir de los signos de los coeficientes se puede señalar que las variables rango de edad con sus opciones de respuesta “entre 770,01 y 1.133 USD (Q2)”, “entre 1.133,01 y 1.603 USD (Q3)”, “entre 1.603,01 y 2.362 USD (Q4)” y “mayor a 2.362 USD (Q5)”; rango de plazo con su opción de respuesta “mayor a 20 años”; subregión geográfica con su opción de respuesta “Costa (incluye Sto. Dom. y Gal.) y Oriente Norte (Nap., Suc., Ore)”; ciclo económico previo al desembolso resumido con su opción de respuesta “(1) desempeño económico bajo crecimiento promedio de largo plazo”; año de concesión; y cesante y ciclo económico por debajo del crecimiento de largo plazo

previo a la mora grave con su opción de respuesta “(1) el cliente sí se encontró simultáneamente en la condición de cesante y en un ciclo económico por debajo del crecimiento de largo plazo previo a la mora grave”; al presentar signo positivo denotaron que su presencia incrementa el riesgo de incurrir en una situación de mora grave.

Por su parte, la variable ciclo económico previo a la mora grave resumido con su opción de respuesta “(1) desempeño económico bajo crecimiento promedio de largo plazo”; al presentar signo negativo indicó que su presencia disminuye el riesgo de incurrir en una mora grave.

n=3570, number of events=695	
	exp(coef)
R_Ingptre 770,01 y 1.133 USD (Q2)	1,458870
R_Ingptre 1.133,01 y 1.603 USD (Q3)	1,469482
R_Ingptre 1.603,01 y 2.362 USD (Q4)	1,708307
R_Ingpmayor a 2.362 USD (Q5)	2,438484
R_Plazomayor a 20 años	1,286739
SubReg2COSTA (incl. STO y GAL.) y ORIENTEN (NAP, SUC,ORE)	1,724306
Ciclo21	2,309123
tyear	3,250016
Cicloef21	0,003348
cesantexciclofl	7,685748
	se(coef)
R_Ingptre 770,01 y 1.133 USD (Q2)	0,111421
R_Ingptre 1.133,01 y 1.603 USD (Q3)	0,111200
R_Ingptre 1.603,01 y 2.362 USD (Q4)	0,118367
R_Ingpmayor a 2.362 USD (Q5)	0,134217
R_Plazomayor a 20 años	0,084211
SubReg2COSTA (incl. STO y GAL.) y ORIENTEN (NAP, SUC,ORE)	0,081800
Ciclo21	0,154043
tyear	0,055452
Cicloef21	0,132724
cesantexciclofl	0,106567

Tabla 5. Modelo de cox – Clientes de créditos hipotecarios que desempeñaron labores en “Empresa unipersonal / construcción / trabajo autónomo, rise y microempresa / sin información” – interpretación del exponencial de los coeficientes.

En una segunda instancia, mediante la interpretación del exponencial de los coeficientes  $\beta$  (*exp(coef)*) se puede manifestar lo siguiente:

Aquellos clientes que se encontraron simultáneamente en la condición de cesante y en una fase del ciclo económico por debajo del crecimiento de largo plazo, fueron 7,69 veces más riesgosos de incurrir en una mora grave que quienes no presentaron estas características.

Así también aquellos clientes que en la variable ciclo económico previo al desembolso se mostraron en la opción "(1) desempeño económico bajo crecimiento promedio de largo plazo" fueron 2,31 veces más riesgosos de incurrir en una mora grave que quienes se mostraron en la opción "(0) desempeño económico sobre crecimiento promedio de largo plazo".

En lo que respecta al año de concesión del crédito se pudo constatar que con cada año que pasó el riesgoso de incurrir en una mora grave se incrementó en 3,25 veces.

Por su parte, también se pudo constatar que a mayor ingreso del cliente existe mayor riesgo de incurrir en una mora grave. Específicamente, en comparación a quienes presentaron ingresos de "hasta 770 USD (Q1)" y se constituyen en la opción base, quienes mostraron ingresos "entre 770,01 y 1.133 USD (Q2)" fueron 1,46 veces más riesgosos de incurrir en una mora grave que aquellos clientes de la opción base. Y de esta manera consecutivamente quienes presentaron ingresos de "entre 1.133,01 y 1.603 USD (Q3)" fueron 1,47 veces más riesgosos, quienes mostraron ingresos de "entre 1.603,01 y 2.362 USD (Q4)" fueron 1,71 veces más riesgosos y finalmente, quienes presentaron ingresos "mayores a 2.362 USD (Q5)" fueron 2,44 veces más riesgosos de incurrir en una mora grave que quienes formaron parte de la opción base.

Adicionalmente, se avizó que aquellos clientes de crédito hipotecario y se localizaron geográficamente en las provincias de la opción "Costa (incluye Sto.Dom. y Gal.) y Oriente Norte (Nap., Suc., Ore.)" fueron 1,72 veces más riesgosos de incurrir en una mora grave que quienes se dispusieron geográficamente en "otras subregiones". Finalmente, entre los exponenciales de los coeficientes que resultaron relevantes se puede interpretar que quienes optaron por un plazo para su crédito hipotecario "mayor a 20 años" fueron 1,29 veces más riesgosos de incurrir en una mora grave que quienes optaron por un plazo "menor a 20 años".

Complementariamente, la función de supervivencia resultante del modelo propuesto denotó que la probabilidad de supervivencia a una mora grave por parte de los clientes del producto préstamo hipotecario que desempeñaron actividades económicas en los sectores "EMPRESA UNI-

PERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN" tendió a contraerse substancialmente a partir del mes 40 (3,33 años).

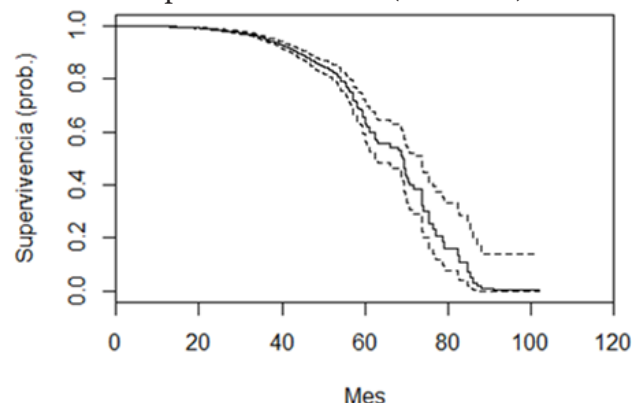


Figura 4. Función de supervivencia resultante de la aplicación del modelo de cox (mora grave > 90 días)

### C. Segmentación de Clientes y Estrategias

1) *Segmentación de Clientes*: En virtud del análisis de supervivencia realizado para el caso de los clientes de créditos hipotecarios de una entidad financiera ecuatoriana se puede definir algunas tipologías de clientes sobre los cuales se pueden definir algunas estrategias específicas. Las tipologías identificadas fueron las siguientes:

- Cliente general de crédito hipotecario.
- Cliente de crédito hipotecario, con ingresos mayores a 1,603 USD de ingreso mensual.
- Cliente de crédito hipotecario, localizado geográficamente en las provincias de la Costa (incluye Santo Domingo y Galápagos) y del Oriente parte norte (Napo, Sucumbíos y Orellana).
- Clientes de Crédito hipotecario que desempeñan labores en los sectores "EMPRESA UNIPERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN".

2) *Estrategias al Momento de la Concesión Crediticia*: Las estrategias que se podrían aplicar al momento de la concesión crediticia en las tipologías de clientes identificadas serían las siguientes:

a) *Cliente general de crédito hipotecario*:

Verificar que se cumplan a cabalidad los requisitos y políticas de crédito al momento de la concesión crediticia, especialmente cuando el entorno económico presenta un desempeño por debajo del crecimiento de largo plazo, dado que la necesidad de generar mayores volúmenes de colocación crediticia hipotecario conduce a otorgar créditos a quienes no necesariamente dan indi-

cios de ser buenos clientes.

*b) Cliente de crédito hipotecario:* con ingresos mayores a 1,603 USD de ingreso mensual: Verificar con mayor intensidad la permanencia en el tiempo de sus ingresos. Identificación clara de la fuente de ingresos. Verificar que el porcentaje comprometido del ingreso mensual por efecto del endeudamiento hipotecario no supere el 40% del ingreso neto.

*c) Cliente de crédito hipotecario:* localizado geográficamente en las provincias de la Costa (incluye Santo Domingo y Galápagos) y del Oriente parte norte (Napo, Sucumbíos y Orellana): Verificar indicios del comportamiento de pago pasado. Peor calificación en los últimos 5 años. Aceptar preferiblemente clientes que hayan mantenido calificaciones de crédito A y B, ya que se identifica un perfil del cliente más riesgoso en estas localidades.

*d) Clientes de Crédito hipotecario:* que desempeñan labores en los sectores "EMPRESA UNIPERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN": Verificar con mayor intensidad la permanencia en el tiempo de sus ingresos. Identificación clara de la fuente de ingresos. Verificar que el porcentaje comprometido del ingreso mensual por efecto del endeudamiento hipotecario no supere el 30% del ingreso neto.

### 3) Estrategias de Gestión de Cobranza:

Las estrategias que se podrían aplicar para optimizar la gestión de cobranza en las tipologías de clientes identificadas serían las siguientes:

*a) Cliente general de crédito hipotecario:* Verificar continuamente, si el cliente incurre en la condición de cesante, particularmente luego del mes 20 y con mayor intensidad a partir del mes 40 luego del otorgamiento del crédito. Recordar mediante mensajes de texto, correo electrónico y llamadas la importancia del repago del crédito. Intensificar la actividad si coyunturalmente la economía atraviesa un desempeño por debajo del crecimiento de largo plazo.

*b) Cliente de crédito hipotecario:* con ingresos mayores a 1,603 USD de ingreso mensual: Verificar continuamente, si el cliente incurre en la condición de cesante, particularmente luego del mes 20 y con mayor intensidad a partir del mes 80 luego del otorgamiento del crédito. Re-

cordar mediante mensajes de texto, correo electrónico y llamadas la importancia del repago del crédito.

*c) Cliente de crédito hipotecario:* localizado geográficamente en las provincias de la Costa (incluye Santo Domingo y Galápagos) y del Oriente parte norte (Napo, Sucumbíos y Orellana): Verificar continuamente, si el cliente incurre en la condición de cesante, particularmente luego del mes 20 y si se localiza en las provincias orientales del norte. Recordar mediante mensajes de texto, correo electrónico y llamadas la importancia del repago del crédito.

*d) Clientes de Crédito hipotecario que desempeñan labores en los sectores "EMPRESA UNIPERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN":* Una vez que se incorpore al modelo de cox desarrollado variables que denoten el comportamiento de pago pasado, modelar específicamente la probabilidad de supervivencia de casos tipo. En virtud de lo anterior se desarrollaran acciones preventivas como recordatorios de pago y otros principalmente entre los meses 20 y 40 de vigencia de los créditos cuando no se ha producido una mora grave de forma previa. La priorización de los casos se da con el siguiente perfil: ingresos mayores a 1,603 USD, localizados en las provincias de la Costa (incluye Santo Domingo y Galápagos) y del Oriente parte norte (Napo, Sucumbíos y Orellana), con plazo de crédito mayor a 20 años, que se encuentren en la condición de cesante, que fueron concedidos durante un ciclo de la economía por debajo del crecimiento de largo plazo.

En virtud de que la variable cesante previo a la mora grave, y más aún, coincidente con la estancía en un ciclo económica por debajo del crecimiento de largo plazo, resultaron incidentes en la probabilidad de incurrir en una mora grave; sería importante desarrollar un modelo de supervivencia específico frente al evento de una mora grave luego de incurrir en el estado cesante, estratificado por la estancia en un ciclo económico por sobre o por debajo del crecimiento de largo plazo. Es decir, estimar el tiempo que tarda un individuo en incurrir en una mora grave luego de caer en el estado cesante e identificar las probabilidades de supervivencia asociadas. Esta actividad permitiría segmentar la base de clientes y

aplicar una estrategia de propuesta anticipada de refinanciamientos y así evitar moras graves.

#### D. Automatización del Análisis de Supervivencia: El caso de una Cartera Hipotecaria Ecuatoriana

1) *Programación en Lenguaje R*: El procedimiento metodológico implementado en el presente trabajo con la finalidad de verificar la utilidad de la aplicación del análisis de supervivencia en la cartera de créditos hipotecarios de una entidad ecuatoriana pudo llevarse a cabo mediante la utilización del lenguaje de programación R y la utilización específica de los software estadísticos R, R estudio y R markdown.

El tratamiento de la información realizado requirió de 5 etapas:

- Preparación de la data.
- Análisis descriptivo de las potenciales variables explicativas.
- Identificación de variables explicativas relacionadas en mayor medida con la variable dependiente.
- Formulación y evaluación del modelo de cox.
- Creación de reportes.

En el apéndice A se incluyó un extracto de la programación correspondiente a la etapa "Formulación y evaluación del modelo de Cox".

#### IV. CONCLUSIONES

En virtud del trabajo realizado se puede afirmar que la hipótesis planteada sí se cumpliría en virtud de las siguientes conclusiones:

- Las experiencias de aplicación del análisis de supervivencia en diferentes ámbitos y particularmente en el ámbito financiero denotaron que la incorporación de este análisis mejora o complementa el realizado con otras técnicas al momento de predecir la probabilidad de incumplimiento (default) como los modelos logísticos, en virtud de que tiene la capacidad de incorporar datos censurados (información perdida en otras técnicas).
- Así también, las citadas experiencias relevaron que la identificación de la temporalidad en la cual se puede incurrir en incumplimiento (default) es un elemento importante de cara a mitigar el riesgo de crédito, para lo cual el típico análisis de matrices de transición es insuficiente. Así

también se vislumbró la necesidad de incorporar a este tipo de modelos variables asociadas al entorno económico y variables microeconómicas específicas de los deudores (8).

- Empíricamente, la aplicación del análisis de supervivencia a una base de datos de clientes del producto préstamo hipotecario en una entidad financiera ecuatoriana, permitió identificar los siguientes comportamientos con respecto a la incursión del cliente en una mora grave (mayor a 90 días):

- La condición de cesante de un cliente, sumada a la estancia temporal en un ciclo económico que se ubica por debajo del crecimiento de largo plazo, se constituye en la condición más relevante identificada de cara a incurrir con mayor probabilidad en una mora grave.

- Los clientes localizados en las provincias de la Costa (incluye Santo Domingo y Galápagos) y del Oriente parte norte (Napo, Sucumbíos y Orellana) son más propensos a caer en una mora grave que aquellos dispuestos en otros lugares.

- A mayor ingreso mensual de los clientes, mayor probabilidad de incurrir en el evento de mora grave.

- Los clientes que desempeñan labores en los sectores "EMPRESA UNIPERSONAL / CONSTRUCCIÓN / TRABAJO AUTÓNOMO, RISE Y MICROEMPRESA / SIN INFORMACIÓN" son más propensos a caer en una mora grave que aquellos que laboran en otros sectores. También permitió identificar que, de manera general, los clientes de crédito hipotecario de la entidad financiera analizada tienden a incrementar su probabilidad de caer en una mora grave a partir del mes 20 y con mayor intensidad a partir del mes 40.

- Los hallazgos especificados permitieron definir potenciales estrategias de concesión y gestión de cobranza para las tipologías de clientes identificados. Es por ello que uno de los aportes importantes del presente trabajo es haber definido y probado un procedimiento metodológico que permite mejorar la mitigación del riesgo de crédito. Específicamente, el procedimiento implementado constó de las siguientes actividades secuenciales: a) definición del evento de interés (mora grave) y con ello definir la variable de interés correspondiente, b) breve descripción de la base de datos a utilizar, c) análisis descriptivo de las



variables consideradas inicialmente relevantes y construcción individual de funciones de supervivencia mediante el método de Kaplan y Meier, d) seleccionar la data y las variables que permitirán definir un modelo de cox estadísticamente significativo y que cumpla con los supuestos de la técnica, e) formulación, validación e interpretación del modelo de cox, y f) segmentación de la base de clientes de la cartera de créditos hipotecarios, a partir del análisis de supervivencia desarrollado y plantear estrategias diferenciadas tanto de concesión como de gestión de cobranzas.

Algunas de las principales recomendaciones derivadas serían las siguientes:

- Con la finalidad de mejorar los resultados predictivos del modelo de Cox desarrollado se recomienda la incorporación de variables que perfilen el comportamiento de crédito pasado del cliente en un horizonte de hasta 5 años.
- Medir periódicamente cada 3 o 6 meses la variable presencia de cesantía con la finalidad de

construir un modelo de cox entendido, en el cual se incorporan variables que dependen del tiempo.

- En virtud de que la variable cesante previo a la mora grave, y más aún, coincidente con la estancia en un ciclo económica por debajo del crecimiento de largo plazo, resultaron incidentes en la probabilidad de incurrir en una mora grave; sería importante para la entidad oferente del crédito hipotecario desarrollar un modelo de supervivencia específico frente al evento de un mora grave luego de incurrir en el estado cesante, estratificado por la estancia en un ciclo económico por sobre o por debajo del crecimiento de largo plazo. Es decir, estimar el tiempo que tarda un individuo en incurrir en una mora grave luego de caer en el estado cesante e identificar las probabilidades de supervivencia asociadas. Esta actividad permitiría segmentar la base de clientes y aplicar una estrategia de propuesta anticipada de refinanciamientos y así evitar moras graves.

#### Referencias

1. González A, López J. Gestión Bancaria. Factores claves en un entorno competitivo. 3era edición. Madrid: McGraw-Hill / Interamericana de España. S.A.U; 2008.
2. Codificación de Resoluciones de la Superintendencia de Banco y la Junta Bancaria. Quito: Superintendencia de Bancos y Seguros; 2015.
3. Comité de Basilea Sobre Supervisión Bancaria. Principios para la administración del riesgo de crédito. Basilea: Bank for International Settlements; 1999.
4. Velandia, N. Establecimiento de un Modelo Logit para la Medición del Riesgo de Incumplimiento en Créditos para una Entidad Financiera del Municipio de Arauca. Departamento de Aucara [Tesis de Post-grado]. Manizales: Universidad Nacional de Colombia; 2013. Recuperado a partir de: <https://repositorio.unal.edu.co/handle/unal/19986>
5. Alfaro-Navarro JL. Control estadístico de la calidad en procesos multivariantes autocorrelacionados. Una aplicación en la industria cuchillera de Albacete [Tesis Doctoral]. Cuenca: Universidad de Castilla-La Mancha; 2005. Recuperado de: <https://www.scribd.com/document/452585681/Tecnica-EWMA-II-pdf>.
6. Pérez-Duque PN. Control estadístico de calidad multivariado, para el monitoreo e identificación de causas de variabilidad en procesos de crédito del sector financiero= Multivariate statistical quality control to monitoring and identifying variability causes in credit process in financial sector [Tesis de Post-grado]. Manizales: Universidad Nacional de Colombia; 2012. Recuperado a partir de: <https://repositorio.unal.edu.co/handle/unal/10931>.
7. Rebas P. Conceptos básicos del análisis de supervivencia. Revista de cirugía española [Internet]. 2005 [Consultado 22 Ene 2018]; 78(4): 222 – 230. Disponible en: <https://www.elsevier.es/es-revista-cirugia-espanola-36-articulo-conceptos-basicos-del-analisis-supervivencia-13079636?referer=buscador>
8. Meira-Machado L. Introduction to Survival Analysis. 1era Edición. Santiago de Compostela: Universidad Santiago de Compostela; 2016.
9. Obuda FY. Analysis of Credit Risk on Bank Loans using Cox's Proportional Hazard Model. Nairobi: Universidad de Nairobi; 2014.

10. Cáceres F, Palacios Y. Análisis de supervivencia como alternativa metodológica para estimar probabilidades de incumplimiento de los deudores de créditos corporativos y a grandes empresas en el Perú. *Revista Industrial Data* [Internet]. 2017 [Consultado 22 Ene 2018]; 20 (1); 07-15. Disponible en: <https://www.redalyc.org/pdf/816/81652135001.pdf>.
11. Cao R, Vilar J, Devia A. “Modelling consumer credit risk via survival analysis”. 1st Edition. Coruña: Universidade da Coruña; 2009.
12. Zhang, A. *Statistical Methods in Credit Risk Modeling* [Tesis Doctoral]. Michigan: The University of Michigan; 2009. Recuperado de: [https://deepblue.lib.umich.edu/bitstream/handle/2027.42/63707/ajzhang\\_1.pdf](https://deepblue.lib.umich.edu/bitstream/handle/2027.42/63707/ajzhang_1.pdf).
13. Lopez A. Análisis de la Relación entre intermediación crediticia y crecimiento económico en Venezuela. 1st Edition. Carácas: Banco Central de Venezuela; 2003.
14. Man R. *Survival analysis in credit scoring: A framework for PD estimation*. 1st Edition Rabobank International – Quantitative Risk Analytics & University of Twente; 2014.
15. Klein JP, Moeschberger ML. *Survival Analysis Techniques for Censored and Truncated Data*. 1st Edition. New York:Springer-Verlag; 1997.
16. Watkins J, Vasnev A, Gerlach R. *Survival Analysis for Credit Scoring: Incidence and Latency*. 1st Edition. Madrid: OME Working Paper Series Madrid; 2019.
17. Herranz Valera J. *Introducción al Análisis de Supervivencia con R*. 1era Edición. Madrid: Instituto IMDEA Alimentación; 2015.
18. Meira-Machado L. *Survival Analysis: Kaplan-Meier estimate*. 1era Edición. Santiago de Compostela: Universidad Santiago de Compostela; 2016.
19. Ayala MA, Borges RE, Colmenares G. Análisis de Supervivencia aplicado a la Banca Comercial Venezolana 1996 – 2004. *Revista Colombiana de Estadística* [Internet]. 2007 [Consultado 22 Ene 2018]; 30(1); 97-113. Disponible en: <https://www.redalyc.org/pdf/899/89930107.pdf>
20. Godoy-Aguilar AM. *Introducción al Análisis de Supervivencia con R*. 1era Edición. México: Universidad Nacional Autónoma de México; 2009.
21. Royston P, Lambert PC. *Flexible parametric survival analysis using Stata: beyond the Cox model* Vol. 347. 1st Edition. College Station: Stata Press; 2011.
22. Almeida E. *Aplicación del Modelo de Supervivencia de Cox al caso de la Banca Ecuatoriana en el período 1996-2008* [Tesis de Grado]. Quito: Escuela Politécnica Nacional; 2011. Recuperado de: <http://bibdigital.epn.edu.ec/handle/15000/4191>.
23. Santabábara J, Rubio E, Martínez T. *Manual de Análisis de Supervivencia: Curvas de supervivencia y Regresión de Cox*. 1era Edición. Zaragoza: Prensas de la Universidad de Zaragoza; 2016.
24. Hernández-Domínguez AM. *Análisis Estadísticos de Datos de Tiempos de Fallo en R*. 1era Edición. Granada: Universidad de Granada; 2010.
25. García-Hinojosa, C. *Estudio de Análisis de Supervivencia*. Madrid: Universidad Autónoma de Madrid, Escuela Politécnica Superior; 2014.