



# **Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

**Facultad de Ciencias Matemáticas**

**Escuela Profesional de Estadística**

## **Análisis de supervivencia para proyectar la deuda en default de la cartera de tarjetas de una entidad financiera**

### **TRABAJO DE SUFICIENCIA PROFESIONAL**

Para optar el Título Profesional de Licenciado en Estadística

#### **AUTOR**

Jashir Alejandro CANLLA LINARES

#### **ASESOR**

Mg. Roberto Carlos FIESTAS FLORES

Lima, Perú

2021



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

Canlla, J. (2021). *Análisis de Supervivencia para proyectar la deuda en default de la cartera de tarjetas de una entidad financiera*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

---

## Referencia bibliográfica

---

Canlla, J. (2021). *Análisis de Supervivencia para proyectar la deuda en default de la cartera de tarjetas de una entidad financiera*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.

---

## Metadatos complementarios

<b>Datos de autor</b>	
Nombres y apellidos	Jashir Alejandro Canlla Linares
Tipo de documento de identidad	DNI
Número de documento de identidad	47383115
URL de ORCID	No aplica
<b>Datos de asesor</b>	
Nombres y apellidos	Roberto Carlos Fiestas Flores
Tipo de documento de identidad	DNI
Número de documento de identidad	16744141
URL de ORCID	<a href="https://orcid.org/0000-0002-5582-0124">https://orcid.org/0000-0002-5582-0124</a>
<b>Datos del jurado</b>	
<b>Presidente del jurado</b>	
Nombres y apellidos	Mg. Carlos Alberto Jaimes Velasquez
Tipo de documento	DNI
Número de documento de identidad	42762905
<b>Miembro del jurado 1</b>	
Nombres y apellidos	Mg. Ricardo Luis Pomalaya Verastegui
Tipo de documento	DNI
Número de documento de identidad	10460674
<b>Datos de investigación</b>	
Línea de investigación	No aplica

Grupo de investigación	No aplica
Agencia de financiamiento	No aplica
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos Latitud: -12.055434 Longitud: -77.082245
Año o rango de años en que se realizó la investigación	2021
URL de disciplinas OCDE	Estadísticas, Probabilidad <a href="https://purl.org/pe-repo/ocde/ford#1.01.03">https://purl.org/pe-repo/ocde/ford#1.01.03</a> Matemáticas aplicadas <a href="https://purl.org/pe-repo/ocde/ford#1.01.02">https://purl.org/pe-repo/ocde/ford#1.01.02</a> Econometría <a href="https://purl.org/pe-repo/ocde/ford#5.02.02">https://purl.org/pe-repo/ocde/ford#5.02.02</a>



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú. Decana de América  
FACULTAD DE CIENCIAS MATEMÁTICAS  
ESCUELA PROFESIONAL DE ESTADÍSTICA

## ACTA DE SUSTENTACIÓN DE TRABAJO DE SUFICIENCIA PROFESIONAL EN LA MODALIDAD VIRTUAL PARA OBTENCIÓN DEL TÍTULO PROFESIONAL DE LICENCIADO EN ESTADÍSTICA

En Lima, siendo las 16:00 horas del domingo 03 de octubre del 2021, se reunieron los docentes designados como Miembros del Jurado del Trabajo de Suficiencia Profesional: Mg. Carlos Alberto Jaimes Velasquez (PRESIDENTE), Mg. Ricardo Luis Pomalaya Verastegui (MIEMBRO) y el Mg. Roberto Carlos Fiestas Flores (MIEMBRO ASESOR), para la sustentación del Trabajo de Suficiencia Profesional titulado: “ANÁLISIS DE SUPERVIVENCIA PARA PROYECTAR LA DEUDA EN DEFAULT DE LA CARTERA DE TARJETAS DE UNA ENTIDAD FINANCIERA”, presentado por el señor **Bachiller Jashir Alejandro Canlla Linares**, para optar el Título Profesional de Licenciado en Estadística.

Luego de la exposición del trabajo de suficiencia, el Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del Jurado Evaluador, el expositor mereció la aprobación **SOBRESALIENTE**, con un calificativo promedio de **DIECISIETE (17)**

A continuación, los miembros del Jurado dan manifiesto que el participante **Bachiller Jashir Alejandro Canlla Linares** en vista de haber aprobado la sustentación del Trabajo de Suficiencia Profesional, será propuesto para que se le otorgue el Título Profesional de Licenciado en Estadística.

Siendo las 16:30 horas se levantó la sesión firmando para constancia la presente Acta.

Mg. Carlos Alberto Jaimes Velasquez  
PRESIDENTE

Mg. Ricardo Luis Pomalaya Verastegui  
MIEMBRO

Mg. Roberto Carlos Fiestas Flores  
MIEMBRO ASESOR

La Vicedecana de la Facultad de Ciencias Matemáticas, Mg. Zoraida Judith Huamán Gutiérrez, certifica virtualmente la participación del Jurado Evaluador, el titulado, el acto de instalación y el inicio, desarrollo y término del acto académico de sustentación, dejando constancia en el acta respectiva.

## RESUMEN

La probabilidad de incumplimiento de los clientes morosos en el sistema financiero es estimada por distintos métodos estadísticos clásicos, dependiendo de la entidad financiera y el tipo de cliente analizado. En este trabajo se desarrolla el análisis de supervivencia para estimar la probabilidad de default de la cartera de tarjetas de una entidad financiera peruana. Además, se realiza un backtest de las probabilidades de default estimadas, para evaluar el modelo con el comportamiento real de la cartera de tarjetas en el año 2020.

Se encontró que al mes 6 de maduración de los desembolsos de la cartera de tarjetas se evidencia la máxima expresión de la probabilidad de incumplimiento para luego en meses posteriores descender lentamente. Finalmente se utilizan las probabilidades estimadas por la función de riesgos para proyectar la deuda en default de los desembolsos de la cartera de tarjetas para el presupuesto del año 2021, utilizando un análisis de cosechas y obteniendo el monto total de default proyectado en una ventana de tiempo de 12 meses.

**Palabras clave:** Análisis de supervivencia, probabilidad de incumplimiento, función de riesgo Hazard, análisis de cosechas, backtest.

## ABSTRACT

The probability of default in the financial system is estimated by different classical statistical methods, depending on the financial institution and the type of customer analyzed. In this paper the survival analysis is developed to estimate the probability of default of the card portfolio of a Peruvian financial institution. In addition, a backtest of the estimated default probabilities is performed, to evaluate the model with the actual behavior of the card portfolio in the year 2020. It was found that month 6 of maturation of the disbursements of the portfolio of cards is evidenced the maximum expression of the probability of default and then in subsequent months slowly decrease. Finally, the probabilities estimated by the risk function are used to project the default debt of the card portfolio disbursements for the 2021 budget, using a crop analysis and obtaining the total amount of default projected in a time window of 12 months.

**Keywords:** Survival analysis probability of default, Hazard risk function, crop analysis, backtest.



## **Tabla de contenido**

I. Introducción .....	4
II. Información del lugar donde se desarrolló la actividad .....	5
III. Descripción de la actividad .....	6
IV. Conclusiones .....	20
V. Recomendaciones .....	21
VI. Bibliografía.....	22

## **I. Introducción**

El negocio de las entidades financieras se basa en recaudar fondos de distintos sectores de la economía para finalmente poder ofrecer productos financieros que ayuden a impulsar el crecimiento económico de un país. En ese sentido las entidades financieras al ofrecer sus productos asumen riesgos de incumplimiento por parte de las contrapartes, este tipo de riesgos se denominan riesgos de crédito en el argot financiero.

La gestión de este tipo de riesgos se aborda de diferente manera de acuerdo al público objetivo de las entidades financieras y una de las herramientas matemáticas comúnmente usadas para medir el riesgo de crédito es el cálculo de la probabilidad de default (PD) de los clientes. Existen varios métodos de cálculo de la PD basados en regresión, matrices de transición, etc. sin embargo en este trabajo se desarrolla el análisis de supervivencia como una alternativa para la estimación de la probabilidad de default (Hernández, 2018).

El análisis de supervivencia es una técnica estadística que se da origen en el siglo XVII y es utilizado en investigaciones clínicas y epidemiológicas; esta herramienta nos permite estudiar el tiempo entre dos eventos cualesquiera permitiéndonos poder aplicarlo en la estimación de la probabilidad de default (Gómez y Cobo, 2016).

En ese sentido el objetivo principal de este trabajo es proyectar el saldo en default de la cartera de tarjetas de una entidad financiera utilizando el análisis de supervivencia.

Se debe tener una adecuada gestión del riesgo de crédito de la entidad financiera para poder garantizar la calidad de clientes y los servicios que se les ofrece, en ese sentido tener una proyección del saldo en default es determinante para que la entidad tenga conocimiento de las posibles pérdidas en las que probablemente incurran y tome acciones preventivas para mantener la solvencia de la misma.

Así el trabajo aborda la problemática en 6 capítulos con mayor detalle:

En el capítulo 2 del presente trabajo se desarrolla una información general de la entidad financiera en donde se realiza el análisis.

En el capítulo 3 se describe el proceso de conocimiento, los objetivos principales del estudio, la problemática y contexto a resolver, el detalle de la metodología a utilizar como una posible solución al problema de estudio y finalmente los principales resultados del estudio.

En el capítulo 4 se desarrollan algunas conclusiones producto del análisis realizado y de las experiencias observadas en la actividad desarrollada en la entidad.

En el capítulo 5 se desarrollan recomendaciones respecto al uso del análisis de supervivencia como una alternativa para estimar la probabilidad de default.

En el capítulo 6 se detallan la bibliografía utilizada para desarrollar el presente trabajo, así como también los anexos del mismo.

## **II. Información del lugar donde se desarrolló la actividad**

- Institución donde se desarrolló la actividad.

La institución donde se desarrolló la actividad pertenece al sistema financiero peruano.

- Periodo de duración de la actividad

La actividad se realizó desde el mes de julio del 2021 al mes de agosto del mismo.

- Finalidad y objetivos de la entidad

La entidad tiene como finalidad prestar los mejores servicios financieros para ayudar a los peruanos a cumplir sus sueños hoy, dentro de los objetivos de la entidad se encuentran la prestación de servicios financieros diversificados a sus clientes con enfoque en productos 100% digitalizados.

- Razón social

Entidad Financiera Peruana

- Dirección postal

No aplica

- Correo electrónico del profesional a cargo

jashir12c@hotmail.com

### III. Descripción de la actividad

- Organización de la actividad (profesional a cargo)
  - Comprensión del negocio  
En esta etapa se realiza una revisión de la literatura respecto a la variable de estudio (default de clientes en el sistema financiero)
  - Comprensión de los datos  
Posterior a la comprensión del negocio se realiza la recolección de los datos de la institución financiera y a su vez una exploración de los mismos para verificar la calidad de la información.
  - Preparación de los datos  
En esta fase se desarrolla todo el proceso de adecuación de los datos para utilizar los métodos de análisis de supervivencia.
  - Modelado  
Se aplica el análisis de supervivencia (función de hazard) a los datos preparados los cuales nos permitieron estimar la probabilidad de incumplimiento de la cartera de tarjetas de la entidad financiera.
  - Evaluación  
Se comparan las curvas de la deuda atrasada real contra la curva de deuda atrasada estimada.
  - Implementación  
Se utiliza el modelo evaluado para predecir la probabilidad de incumplimiento de los clientes, además se realiza un análisis de cosechas para proyectar el saldo en default de la cartera de tarjetas.

- Finalidad y objetivos de la actividad

#### Finalidad

Proyectar el saldo en default de los clientes de la cartera de tarjetas de una entidad financiera.

#### Objetivos Específicos

- Estimar la función de supervivencia de los clientes que entran en default de la cartera de tarjetas de la entidad financiera.

- Estimar la función de densidad de los clientes que entran en default de la cartera de tarjetas de la entidad financiera.
  - Estimar la función de Riesgo (Función de Hazard) de los clientes que entran en default de la cartera de tarjetas de la entidad financiera.
  - Desarrollar un análisis Backtest para validar las probabilidades de default obtenidas del modelo de Hazard.
  - Desarrollar un análisis de cosechas para proyectar el capital en default de la cartera de tarjetas de la entidad financiera.
- Problemática

El presente trabajo analizó una serie de antecedentes que aborda la problemática en estudio y se presentan a continuación:

Joubert, Verster y Raubenheimer (2018) proponen utilizar el análisis de supervivencia como una mejora al método más tradicional de regresión logística para estimar la probabilidad de default en uno de los más grandes bancos retail del sur africano. Se evaluaron los modelos con las métricas del error cuadrático medio, el sesgo y la varianza, encontrando que el modelo de regresión logística obtenía un MSE de 0.163 contra el modelo de supervivencia con un MSE de 0.009, en la varianza de los modelos se encontró que la regresión logística obtenía un valor de 0.162 mientras para el modelo de supervivencia resultaba con una varianza inferior de 0.008, además se observó que el sesgo del modelo de supervivencia era inferior con -0.037 contra un valor de -0.028 para la regresión logística. Demostraron así un mayor poder predictivo del modelo de supervivencia, además concluyen que el análisis de supervivencia tiene más ventajas como permitir la censura y poder obtener predicciones en distintos periodos.

Uriarte y Agüero (2017) plantean el análisis de supervivencia como una opción metodológica mejorada y ventajosa respecto al método tradicional de matrices de transición para estimar probabilidades de no pago en los clientes de grandes empresas y corporativos. Además, identifican nuevos insights que explican el riesgo de incumplimiento de dichos segmentos de

clientes gracias a que la metodología de análisis de supervivencia lo permite. El horizonte de tiempo comprende a los clientes entre los años 2003 y 2012. Concluyen que las variables sector de la economía al que pertenecen, nivel de deuda, rentabilidad, magnitud de la empresa, ratio de garantías, agrupación e índice de confianza empresarial que explican el atraso en el cumplimiento de sus obligaciones financieras, y por lo tanto se recomienda hacer un seguimiento de estas variables con el propósito de mejorar el perfil de este tipo de clientes en las entidades.

Rychnovský (2018) sostiene que muchas investigaciones demuestran que los resultados obtenidos en la estimación de la probabilidad de default por los modelos de regresión logística y los modelos de supervivencia son muy similares, sin embargo, plantea una nueva forma de comparar estos modelos a través de una muestra ex-ante de validación en la cual los modelos de análisis de supervivencia serían mejores. Considera a los coeficientes Gini y Lift para poder evaluar el poder predictivo de cada modelo en una data real de uno de los más grandes bancos de República Checa en el horizonte de tiempo de 2002-2006. Concluye que los modelos tienen un comportamiento similar en la muestra aleatoria simple con coeficientes Gini de 0.39 y 0.40 para el modelo logístico y de Supervivencia respectivamente Sin embargo en la muestra ex – ante de validación se observa que el modelo de Cox (Análisis de Supervivencia) supera al modelo logístico obteniendo valores Gini 0.43 y 0.38 respectivamente. Finalmente concluye que su investigación aporta al debate en el ámbito del modelamiento matemático en finanzas y muestra este tipo de nuevos criterios para evaluar modelos predictivos a través del tiempo.

Dirick, Claeskens y Baesens (2017) consideran que los estudios de estimación de probabilidad de default entre los años 1992 y 2015 utilizan diversos métodos de análisis de supervivencia con diferentes métricas de evaluación, sin embargo, hasta ese momento no había investigaciones que compararan todos estos tipos de modelos en diferentes datasets. Así en este trabajo comparan métodos de análisis de supervivencia tradicionales (Accelerated failure time models y Cox proportional hazards regression models) y modelos con variantes nuevas como Mixture cure models para uno y varios eventos que se introdujeron en el 2017 en el ámbito del riesgo de crédito. Se utilizaron 10 datasets de instituciones financieras de Bélgica y Reino Unido, en las cuales se utilizaron métricas estadísticas de evaluación (AUC y las

predicciones de créditos en default) y métricas económicas (teoría de la anualidad). Concluyen que los modelos de Cox PH con splines penalizados ( $\overline{AUC}$ : 3.3, MSE: 2.9, MAE: 2.8) son los que obtienen mejores métricas de evaluación respecto a los demás modelos evaluados.

Gupta (2017) identifica predictores claves que permitan explicar la probabilidad de default para 859 compañías de la India evaluadas en el rango de tiempo de los años 2000 y 2015, información recolectada de las bases de datos del centro de seguimiento de la economía india (CMIE). Las variables evaluadas abordan la rentabilidad, solvencia, liquidez, variables macroeconómicas, variables internas de las compañías y variables del mercado indio. Utiliza el análisis de supervivencia (semi-parametric Cox proportional hazard model) para encontrar la relación entre estas variables y la probabilidad de default de las compañías en estudio. El modelo arroja un AUC de 81,9% denotando un buen poder de predicción, además se concluye que las variables rendimiento del capital empleado, rendimiento del patrimonio neto, ratio de cobertura de intereses, volatilidad del tipo de cambio, tasa de crecimiento, índice bursátil, participación bursátil (%) y porcentaje de acciones pignoradas son significativas para predecir la probabilidad de default.

- Metodología, Procedimientos

Enfoque: Cuantitativo, ya que abordamos la problemática basada en la recolección de datos como insumos para la utilización de herramientas matemáticas, estadísticas e informáticas.

Diseño: No experimental, ya que la información recolectada es obtenida con la observación del fenómeno dentro de su determinado contexto.

Nivel: Explicativo, pues estudiamos la relación que existe entre el tiempo que transcurre hasta llegar a un estado de incumplimiento de los clientes con el nivel de saldo mes a mes de los mismos.

Población: Clientes de la cartera de tarjetas de crédito de una entidad financiera peruana en el periodo de enero 2019 a diciembre 2020.

Se definen conceptos para entender la importancia de la gestión del riesgo de crédito y por consiguiente la importancia de la proyección del saldo en default a través de la estimación de la probabilidad de default en una entidad financiera de acuerdo a la regulación local de la Superintendencia de Banca, Seguros y AFP (SBS).

- Patrimonio Efectivo

El patrimonio efectivo de una entidad financiera es el importe extra-contable que es determinado sumando el patrimonio básico y el patrimonio suplementario del mismo. (Ley N°26706 de 2019 [Superintendencia de Banca, Seguros y AFP]. texto concordado de la ley general del sistema financiero y del sistema de seguros de la superintendencia de banca y seguros.)

- Requerimientos de Patrimonio Efectivo

La SBS proporciona algunos lineamientos generales para medir los requerimientos que exige la misma por conceptos de riesgo de crédito, riesgo de mercado y riesgo operacional de las entidades financieras. En el particular para riesgo de crédito establece un método estándar y la opción de emplear modelos internos para el cálculo del mismo. (Ley N°26706 de 2019 [Superintendencia de Banca, Seguros y AFP]. texto concordado de la ley general del sistema financiero y del sistema de seguros de la superintendencia de banca y seguros.)

- Riesgo de Crédito

La SBS hace referencia al riesgo de crédito como la exposición al incumplimiento de las condiciones de su contrato financiero de la contra-parte (cliente). (Ley N°26706 de 2019 [Superintendencia de Banca, Seguros y AFP]. texto concordado



de la ley general del sistema financiero y del sistema de seguros de la superintendencia de banca y seguros.)

- Clientes en Default

Son las contra-partes que incurren en incumplimiento de sus obligaciones financieras y permanecen en este estado por más de 90 días para efectos de esta investigación.

- Probabilidad de Default

Es la cuantificación del riesgo de incumplimiento de la contra-parte mediante la teoría de probabilidades y modelos estadísticos.

- Saldo en Default

Monto capital de deuda de la contra-parte en estado de default para la entidad financiera.

## Análisis de Supervivencia

El análisis de supervivencia es una técnica estadística que tiene su origen en estudios médicos ya desde el siglo XVII, sin embargo, con el transcurrir de los años se ha adaptado la técnica para resolver muchas problemáticas de distinta naturaleza, por ejemplo, en el análisis de *churn* (fuga de clientes).

El análisis de supervivencia estudia el tiempo que transcurre hasta que se observa el evento de interés que puede ser, por ejemplo, el tiempo hasta que un cliente cancele el servicio, el tiempo hasta que un paciente presente algún síntoma adverso, etc. En ese sentido al estudiar este fenómeno se presentan casos en donde no se observa el evento de interés los cuales son denominados observaciones censuradas. Esta información censurada es muy

importante a la hora de modelar el tiempo de supervivencia, ya que nos proporciona información mucho más real al agregar este componente al modelo, así el análisis de supervivencia es muy ventajoso frente a otras técnicas estadísticas como el análisis de regresión.

Entonces en el análisis de supervivencia para cada observación estudiaremos el tiempo de supervivencia  $T$  o en su defecto el tiempo de Censura  $C$ . Así se define la variable aleatoria:

$$Y = \min(T, C). \quad (1)$$

También se define el indicador de estado  $\delta$ :

$$\delta = \begin{cases} 1 & T \leq C \\ 0 & T > C. \end{cases}$$

Donde  $\delta=1$  indica que se observa el tiempo de supervivencia y  $\delta=0$  si observamos el tiempo de censura.

### Curva de Supervivencia Kaplan-Meier

La función de supervivencia se define como:

$$S(t) = \Pr (T > t) \quad (2)$$

Esta función decreciente cuantifica la probabilidad de supervivencia pasado un tiempo  $t$ . Cuanto mayor sea  $S(t)$  menor es la probabilidad que el tiempo de supervivencia sea menor a  $t$ .

Estimar esta función no es tarea sencilla por la existencia de los datos censurados en este tipo de problemas, sin embargo, gracias a la ley de probabilidad total se puede estimar la curva como:

$$\hat{S}(d_k) = \prod_{j=1}^k \left( \frac{r_j - q_j}{r_j} \right) \quad (3)$$

Donde  $r_j$  es el número de observaciones en riesgo hasta antes del tiempo de muerte  $d_j$  y  $q_j$  el número de observaciones que mueren en el tiempo  $d_j$ .

Para tiempos  $t$  entre  $d_k$  y  $d_{k+1}$  se fija que  $S(t) = S(d_k)$  por ende la curva de supervivencia Kaplan-Meier presenta una forma escalonada.

### Función de Riesgo

La función de riesgo conocida como la intensidad de mortalidad en el análisis de supervivencia se define como:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T \leq t + \Delta t | T > t)}{\Delta t} \quad (4)$$

donde  $T$  es el tiempo de supervivencia. Esta función representa el ratio de mortalidad, sin embargo, al tomar el límite cercano a 0 esta función se aproxima a la función de densidad para  $T$  dado  $T > t$  de la siguiente manera:

$$h(t) \approx \frac{\Pr(t < T \leq t + \Delta t | T > t)}{\Delta t} \quad (5)$$

para algún pequeño arbitrario  $\Delta t$ .

También gracias a la probabilidad condicional que se puede expresar como  $\Pr(A|B) = \Pr(A \cap B) / \Pr(B)$  entonces se tiene la relación entre la función de riesgo  $h(t)$ , la función de supervivencia  $S(t)$  y la función de densidad  $f(t)$  de la siguiente manera:

$$h(t) = \frac{f(t)}{S(t)}$$

donde  $f(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t < T \leq t + \Delta t)}{\Delta t}$  es el ratio de mortalidad en el tiempo  $t$ .

- Resultados de la actividad

Se han evaluado 22,993 créditos de la cartera de tarjetas en un horizonte de 12 meses comprendidos entre enero del 2020 y diciembre 2020.

Se recolectó la información concerniente al total de créditos por mes, los créditos en default, es decir aquellos créditos que sobrepasan los 90 días de mora con la entidad financiera y finalmente los créditos que se cancelan con frecuencia mensual.

**Tabla 1**  
**Créditos de la cartera Tarjetas Ene-20 - Dic20**

Mes	Número de Créditos	Créditos Mora > 90 días	Créditos Cancelados
ene-20	22,993	-	8
feb-20	22,985	-	310
mar-20	22,675	-	625
abr-20	22,050	165	892
may-20	20,993	1,090	1,057
jun-20	18,846	1,098	1,310
jul-20	16,438	250	2,187
ago-20	14,001	219	3,720
sep-20	10,062	135	4,650
oct-20	5,277	112	2,312
nov-20	2,853	60	765
dic-20	2,028	45	321

Fuente: Elaboración propia.

Observamos en la tabla 1 que a partir del tercer mes de maduración de los desembolsos se comienza a evidenciar los clientes que entran en default, el pico más alto de los clientes en default se da en junio 2020 para luego caer a niveles de tan solo 45 clientes en default.

Para aplicar el análisis de supervivencia necesitamos calcular las probabilidades de caer en default por cada tramo de tiempo en evaluación considerando la información censurada (créditos cancelados).

Los créditos en riesgos son los créditos expuestos al default en cada tramo de tiempo considerando la información de los créditos cancelados, luego la probabilidad de caer en default se calcula de la división entre los créditos que incumplen sobre el total de créditos en riesgo efectivo. Por complemento se calcula probabilidad de no caer en default. Los resultados se observan en la tabla 2.

*Tabla 2*  
*Probabilidad de Default por tramo de tiempo*

tramo de tiempo	Créditos en Riesgo	Prob. malo	Prob. bueno
1	22,989	-	1.000
2	22,830	-	1.000
3	22,363	-	1.000
4	21,604	0.008	0.992
5	20,465	0.053	0.947
6	18,191	0.060	0.940
7	15,345	0.016	0.984
8	12,141	0.018	0.982
9	7,737	0.017	0.983
10	4,121	0.027	0.973
11	2,471	0.024	0.976
12	1,868	0.024	0.976

Fuente: Elaboración propia.

En la figura 1 se observa la función de densidad de caer en default (a:  $f(t)$ ), el comportamiento en 12 meses de esta función toma su pico más alto al mes 6 de observación para luego descender y mantenerse estable.

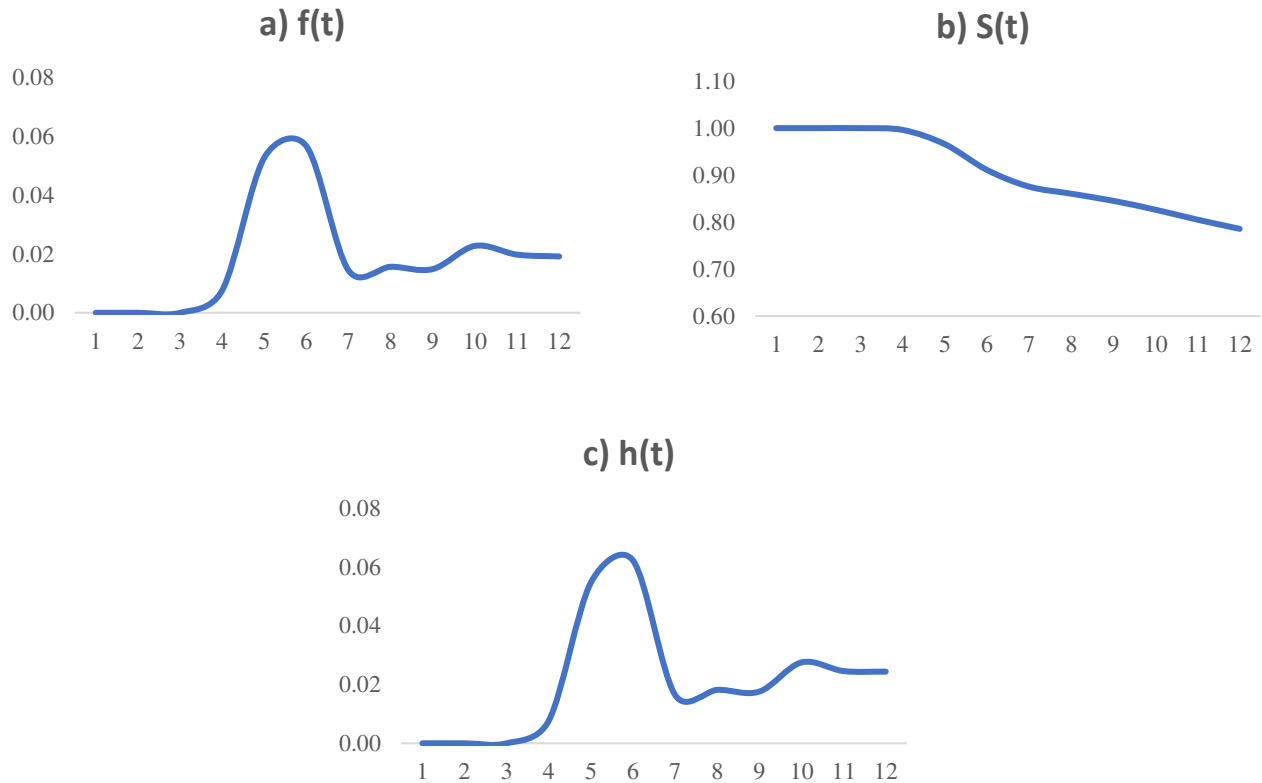
La figura b):  $S(t)$  nos muestra el comportamiento de la curva de supervivencia la cual muestra los niveles más alto de probabilidad en los primeros meses de observación para luego con el transcurrir del tiempo disminuya.

La figura c):  $h(t)$ , muestra la función de riesgo la cual resulta del cociente entre la función de densidad  $f(t)$  y la función de supervivencia  $S(t)$ . El comportamiento de esta curva es

similar a la función de densidad  $f(t)$  ya que recoge el comportamiento de las operaciones evaluadas.

Figura 1

Función de Densidad  $f(t)$ , Curva de Supervivencia  $S(t)$  y Función de Riesgo  $h(t)$



Fuente: Elaboración propia.

Luego de estimar la función de riesgo evaluamos si este cálculo se acerca a la realidad del comportamiento de los clientes evaluados. Para esta tarea se observa el stock de la cartera de tarjetas de diciembre 2020 y se analizan los desembolsos de enero 2020 a diciembre 2020 que pertenecen al stock en mención.

En la tabla 3 se observa que los desembolsos de julio 2020 tienen la mayor proporción de deuda en default, cabe apuntar que estos desembolsos tienen 5 meses de maduración a diciembre 2020.

El trimestre más antiguo con desembolsos de enero 2020 a marzo 2020, tienen en promedio 11.7% de saldo en default respecto de su total de colación al cierre de diciembre 2020.

Cabe mencionar que los créditos refinanciados y judiciales son excluidos del análisis de estimación de las curvas de supervivencia, así como la evaluación del backtest.

*Tabla 3*

*Desembolsos del Stock de Tarjetas a diciembre 2020*

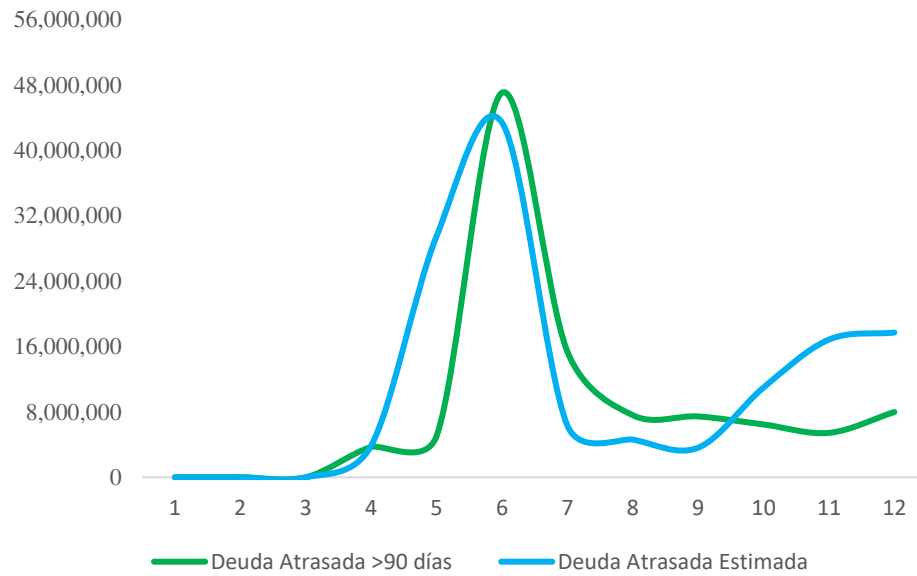
<b>STOCK</b>	<b>MES_DES</b>	<b>Deuda Atrasada &gt;90 días</b>	<b>Desembolso S/</b>
202012	202012	0	72,933,352
202012	202011	0	37,146,108
202012	202010	0	39,684,402
202012	202009	3,713,389	50,390,063
202012	202008	5,018,174	53,866,705
202012	202007	47,047,385	69,768,196
202012	202006	15,403,258	38,646,716
202012	202005	7,605,871	25,336,198
202012	202004	7,450,378	20,581,558
202012	202003	6,466,730	39,860,813
202012	202002	5,442,323	68,548,261
202012	202001	7,989,752	72,627,046

Fuente: Elaboración propia.

En la figura 2 se muestra el resultado del backtest, es decir, observamos la deuda con mayor a 90 días real al cierre de diciembre 2020 contra la curva de la deuda atrasada mayor a 90 días estimada por la función de riesgo  $h(t)$  calculada en la figura 1. Se observa que la realidad como la estimación se comportan de manera similar obteniendo el pico más alto a la mitad de maduración de los desembolsos de créditos de tarjetas.

Figura 2

Backtest al cierre de diciembre 2020



Por último, podemos utilizar las probabilidades estimadas de la función de riesgos  $h(t)$  para realizar un forecast utilizando los desembolsos esperados para el año 2021.

Tabla 4

Desembolsos esperados para la gestión 2021

Cosecha	Desembolso
202101	50,000,000
202102	55,000,000
202103	60,000,000
202104	65,000,000
202105	70,000,000
202106	75,000,000
202107	80,000,000
202108	85,000,000
202109	90,000,000
202110	95,000,000
202111	100,000,000
202112	105,000,000

Utilizamos un análisis de tipo cosecha para los desembolsos esperados para el 2021 y la matriz de probabilidades (tabla 5) de incumplimiento estimadas por la función de riesgo.



Tabla 5

Matriz de Probabilidades de Default

202101	202102	202103	202104	202105	202106	202107	202108	202109	202110	202111	202112
-	-	-	0.01	0.05	0.06	0.02	0.02	0.02	0.03	0.02	0.02
	-	-	-	0.01	0.05	0.06	0.02	0.02	0.02	0.03	0.02
		-	-	-	0.01	0.05	0.06	0.02	0.02	0.02	0.03
			-	-	-	0.01	0.05	0.06	0.02	0.02	0.02
				-	-	-	0.01	0.05	0.06	0.02	0.02
					-	-	-	0.01	0.05	0.06	0.02
						-	-	-	0.01	0.05	0.06
							-	-	-	0.01	0.05
								-	-	-	0.01
									-	-	-
										-	-
											-

La tabla 6 muestra la deuda atrasada por mes de maduración y mes de desembolso, en promedio se espera una deuda atrasada de 7.2 MM de soles por mes. Al cierre de diciembre del 2021 se espera una deuda atrasada total de 18.2MM. Considerar que se asume un crecimiento lineal de los desembolsos para el año 2021.

Tabla 6

Proyección de la deuda atrasada para el año 2021.

202101	202102	202103	202104	202105	202106	202107	202108	202109	202110	202111	202112
-	-	-	383,338	2,736,012	3,111,892	821,315	910,111	880,110	1,377,614	1,229,256	1,219,512
	-	-	-	421,671	3,009,614	3,423,081	903,446	1,001,122	968,120	1,515,375	1,352,182
		-	-	-	460,005	3,283,215	3,734,270	985,578	1,092,133	1,056,131	1,653,137
			-	-	-	498,339	3,556,816	4,045,460	1,067,709	1,183,144	1,144,142
				-	-	-	536,673	3,830,417	4,356,649	1,149,841	1,274,155
					-	-	-	575,006	4,104,019	4,667,838	1,231,972
						-	-	-	613,340	4,377,620	4,979,027
							-	-	-	651,674	4,651,221
								-	-	-	690,008
									-	-	-
										-	-
											-
-	-	-	383,338	3,157,684	6,581,511	8,025,950	9,641,317	11,317,693	13,579,584	15,830,880	18,195,357

#### IV. Conclusiones

El análisis de supervivencia aplicada a la banca, en específico en temas de riesgo de crédito, otorga muchas ventajas para generar información relevante en la toma de decisiones y planeamiento de la institución de una forma rápida, concisa, específica y clara.

El presente trabajo nos permite concluir lo siguiente:

- La curva de supervivencia estimada de forma no paramétrica nos da a entender que la probabilidad de no caer en default (sobrevivir al evento) en el tiempo evaluado tiene un pico máximo de 99.6% al mes 4 de maduración y luego esta probabilidad decae por el mismo deterioro de los desembolsos.
- La función de densidad nos muestra un comportamiento cuasi normal de la probabilidad de caer en default, llegando en promedio a obtener 1.8% de probabilidad, la cual se encuentra dentro del apetito al riesgo de la entidad financiera.
- La función de riesgo nos da entender la probabilidad de default en los clientes de tarjetas de crédito evaluados dado que se evidencia una probabilidad de supervivencia en una determinada unidad de tiempo, así la máxima probabilidad de default se da al mes 6 de maduración de los desembolsos con un de 6.2%, por otro lado, en promedio se obtiene probabilidad 2.5% a 12 meses de maduración.
- La proyección del saldo en default para el año 2021 de los desembolsos estimados por la institución financiera ascienden a 18.2MM, esta cantidad representa el 1.96% de total de desembolsos de un 1 año comercial.

## **V. Recomendaciones**

Con este trabajo se evidencia una herramienta práctica y ventajosa para la institución en el sentido de proyectar los saldos en default para una gestión anticipada del riesgo de crédito. Se recomienda actualizar las probabilidades de default de manera mensual para observar el comportamiento de los desembolsos de la cartera de tarjetas con el fin de actualizar los pronósticos y asegurar la correcta gestión del portafolio para no exceder los apetitos al riesgo de la entidad financiera.

Además, se recomienda analizar otros tipos de créditos como el de no retail para poder observar su comportamiento y la correcta adecuación del análisis de supervivencia para proyectar la deuda en default.

## VI. Bibliografía

- Dirick, L., Claeskens, G. & Baesens, B. (2017). *Time to default in credit scoring using survival analysis: a benchmark study*. Journal of the Operational Research Society.
- Gareth, J., Witten, D., Hastie, T. & Tibshirani, R. (2021). *An Introduction to Statistical Learning - with Applications in R*. Springer Science+Business Media, LLC, part of Springer Nature 2021.
- Gómez, G. & Cobo, E. (2016). *Hablemos de Análisis de Supervivencia*. Elsevir.  
<https://www.elsevier.es/index.php?p=revista&pRevista=pdfsimple&pii=70000203&r=36>
- Joubert, M., Verster, T. & Raubenheime, H. (2018). *Making use of survival analysis to indirectly model loss given default*. ORiON.
- Lee, E. & Wenyu Wang, J. (2017). *Statistical Methods for Survival Data Analysis*. John Wiley & Sons, Inc.
- Moore, D. (2017). *Applied Survival Analysis Using R*. Springer International Publishing Switzerland 2016.
- Rychnovský, M. (2018). *Survival analysis as a tool for better probability of default prediction*. Acta Oeconomica Pragensia.
- Texto Concordado De La Ley General Del Sistema Financiero Y Del Sistema De Seguros Y Orgánica De La Superintendencia De Banca Y Seguros, Ley N° 26702, del 9 de diciembre de 1996.
- Uriarte Cáceres, F. & Agüero Palacios, Y. (2017). *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.
- Vandana, G. (2017). *A Survival Approach to Prediction of Default Drivers for Indian Listed Companies*. Theoretical Economics Letters.