

Análisis de Datos Funcionales Aplicado al Riesgo Financiero: un Caso de Estudio en Cooperativas de Ahorro y Crédito Ecuatorianas

Functional Data Analysis Applied to Financial Risk: a Case Study in Ecuadorian Credit Unions

Sergio Castillo-Páez, Miguel Flores, Giovanni Herrera-Enríquez

Resumen—Existe una gran variedad de herramientas estadísticas desarrolladas para la toma de decisiones en el contexto empresarial. El análisis de datos funcionales es una rama de estudio de creciente importancia en los últimos años, y en el presente trabajo se proponen algunas de sus técnicas aplicadas en el contexto del análisis de riesgo financiero. Específicamente, se presentan resultados de análisis funcional exploratorio, identificación de datos atípicos y la construcción de modelos de clasificación supervisada en base a la clasificación del riesgo de las cooperativas de ahorro y crédito, sujetas al control de la Superintendencia de Bancos del Ecuador, tomando como variables funcionales al índice de morosidad y el margen neto de utilidad, en el período julio 2011 a diciembre 2012.

Palabras clave—Datos Funcionales, riesgo financiero, clasificación supervisada.

Abstract—There is a wide variety of statistical tools developed for making decisions in the business context. The analysis of functional data is a area of study of growing importance in the last years. In the present paper some of its applied techniques are proposed to make analysis of financial risk. Specifically, results from exploratory functional analysis, identification of atypical data and the construction of supervised classification models based on the risk classification of credit unions, subject to the control of the Superintendency of Banks of Ecuador, taking as functional variables the NPL ratio and the net profit margin in the period July 2011 to December 2012.

Index Terms—Functional Data, financial risk, supervised classification.

I. INTRODUCCIÓN

Las diferentes crisis que en la última década han afectado al sistema financiero mundial, se deben en gran medida a errores secuenciales o sistemáticos que no fueron

advertidos en los procesos de control, sumado entre otros factores, a la proliferación de instrumentos sofisticados, como los derivados financieros que superaron la capacidad de las instituciones [10], para advertir cambios que provocarían el colapso económico, no solo de grandes corporaciones como LehmanBrothers, sino también de países, siendo el caso más relevante la Eurozona [20].

Una de las actividades vitales de las instituciones financieras, es el control del riesgos financieros, siendo el riesgo de crédito probablemente el de mayor importancia al que se enfrentan las instituciones financieras [4, p. 16]. El riesgo puede entenderse como la probabilidad de enfrentar pérdidas, sin embargo, según [10], propone entender el riesgo como la probabilidad de observar rendimientos distintos a los que se tienen proyectados. Al ser una probabilidad, la gestión del riesgo se vuelve una necesidad ineludible, en toda institución financiera.

Así, el análisis del riesgo financiero es un tema de actualidad y está ocupando el desarrollo de la actividad científica, la misma que mantiene como premisa, que el adecuado funcionamiento del sistema financiero depende de la confianza de los agentes, en un grado mayor que en otros sectores de la economía [20]. Es importante considerar, que a más de las variables relacionadas directamente con el riesgo financiero, la rentabilidad de las empresas en este sector, depende de los niveles de estrés macroeconómico y financiero [9], por lo que es necesario incorporar aspectos regulatorios que protejan su buen desempeño.

Dentro de este contexto, la medición y gestión de riesgos, son actividades de primordial importancia para las instituciones financieras a nivel mundial, así como para los organismos de supervisión [8]. Por tal motivo, son variadas las técnicas y herramientas disponibles para realizar una adecuada gestión y control de los distintos tipos de riesgo financiero: de mercado, operativo, crediticio, legal, entre otros [4]. Algunas de estas técnicas combinan información financiera con métodos matemáticos y estadísticos, como el análisis discriminante [1], redes neuronales [21], series de tiempo, árboles de decisión, entre otros [17].

Usualmente, la información financiera está compuesta por distintas variables medidas a lo largo del tiempo. Sin embargo,

Article history:

Received 15 March 2017

Accepted 11 May 2017

S. Castillo-Páez is a professor at the Departamento de Ciencias Exactas of the Universidad de las Fuerzas Armadas del Ecuador ESPE, Sangolquí, Pichincha, Ecuador (e-mail: sacastillo@espe.edu.ec)

M. Flores, is a professor at the Departamento de Matemática, Escuela Politécnica Nacional, 17012759 Ecuador (e-mail: miguel.flores@epn.edu.ec)

G. Herrera-Enríquez is a professor at the Departamento de Ciencias Económicas, Administrativas y de Comercio of the Universidad de las Fuerzas Armadas del Ecuador ESPE, Sangolquí, Pichincha, Ecuador (e-mail: gpherrera@espe.edu.ec)

estas series de tiempo se suelen analizar con herramientas estadísticas basadas en ciertos supuestos (p.e. estacionariedad, periodicidad, etc., en las series temporales) que no necesariamente reflejan la volatilidad propia de este tipo de variables [3]. Por otro lado, si esta información se expresa como funciones definidas sobre un dominio temporal continuo, entonces se vuelve necesario contar con técnicas apropiadas para su estudio. El análisis de datos funcionales (FDA), es una especialidad estadística en desarrollo, con el objetivo de construir herramientas en las cuales las variables de estudio corresponden a funciones definidas sobre un dominio real [13]. Sus aplicaciones abarcan un amplio espectro, desde el estudio en el uso de suelos [2], medicina [16], entre otros. Las bondades de estos métodos, han sido aplicados también en el contexto económico y financiero, por ejemplo, en el análisis del comportamiento del índice de bienes perecibles [14], sin embargo su aplicación en el contexto del análisis del riesgo financiero no ha sido estudiado de forma exhaustiva.

El presente trabajo, se presenta una aproximación inicial del análisis de datos funcionales para estudiar ciertos ratios financieros de las cooperativas de ahorro y crédito del Ecuador. Si bien existen numerosas herramientas para el análisis de riesgo financiero (que se mencionaron anteriormente), el principal objetivo es aplicar herramientas básicas del análisis funcional en este contexto para verificar su utilidad potencial y la consistencia teórica y práctica de los resultados obtenidos a partir de su aplicación.

Este documento está dividido en los siguientes apartados: En la sección 2 se detallan algunas definiciones del análisis de datos funcionales, y se describen de manera general el conjunto de datos utilizados. Con esta información, se realiza un estudio de datos funcionales cuyos resultados se exponen en la sección 3. En este apartado se realiza la representación funcional de los índices financieros, haciendo énfasis en el análisis exploratorio y la identificación de curvas atípicas y bandas de confianza. Luego en la sección 4 se presentan algunos modelos de clasificación supervisada, usando al índice de morosidad y al margen de utilidad como variables clasificadoras. Las conclusiones obtenidas se presentan en la sección 5.

II. DEFINICIONES Y DATOS UTILIZADOS

Tomando en cuenta que el alcance del presente trabajo es un estudio exploratorio más que exhaustivo de los métodos de la FDA, se recurrirán a ciertas nociones y definiciones básicas las cuales pueden consultarse de forma detallada en los libros de Ramsay y Silverman [14] y Ramsay y Silverman [15]. El libro de Ferraty y Vieu [7] es otra referencia importante que incorpora enfoques no paramétricos, así como el uso de otras herramientas teóricas como Semi-normas y probabilidades de bola pequeña que nos permiten hacer frente a normas o espacios métricos.

Variable Funcional Sea \mathcal{X} una variable aleatoria se dice que es funcional si toma valores en el espacio funcional \mathcal{E} , el cual puede ser normado o seminormado. Por lo general, los datos funcionales están discretizados en un conjunto de puntos $\{t_j\}$, $j = 1, \dots, d$, con $d > 0$, no necesariamente equidistantes, Por simplicidad, utilizaremos la notación $\mathcal{X} = \mathcal{X}(t)$.

Un conjunto de datos funcionales $\{x_1, \dots, x_n\}$ es la observación de n variables funcionales $\mathcal{X}_1, \dots, \mathcal{X}_n$ idénticamente distribuidas

Al trabajar con datos funcionales, el paso inicial es determinar el espacio funcional (con su métrica respectiva) en el cual se van a representar los datos. Un espacio funcional comúnmente utilizado es el espacio \mathbb{R}^n conjuntamente con su métrica euclídea habitual. Cuando el área entre curvas puede proporcionar información relevante acerca del comportamiento de los datos, es usual recurrir al espacio funcional \mathcal{L}_2 (el espacio de Hilbert de todas las funciones integrables cuadradas sobre un cierto intervalo). Para esto, suponiendo un dominio $T = [a, b] \subset \mathbb{R}$ se define:

$$\mathcal{L}^2 = \left\{ \mathcal{X} : T \Rightarrow \mathbb{R}, \text{ tal que } \int_T \mathcal{X}^2(t) dt < \infty \right\}.$$

Asimismo, en ciertos casos los datos pueden requerir algún tipo de transformación, mediante su representación en una base funcional específica. Un conjunto de funciones conocidas e independientes $\{\phi_k\}_{k \in \mathbb{N}}$ es una base de un espacio funcional \mathcal{E} , si cualquier función (dato funcional) $\mathcal{X} \in \mathcal{E}$ puede aproximarse por una combinación lineal finita tal que:

$$\mathcal{X}(t) \approx \sum_{k=1}^K c_k \phi_k(t), \quad (1)$$

donde c_k son constantes. Las bases funcionales más utilizadas son la base Natural, B spline y la base de Fourier.

Posteriormente, se recurre al análisis exploratorio funcional en el cual se calculan varios estimadores, los cuales dependen de la métrica (o distancia) d definida sobre el espacio funcional \mathcal{E} . De manera general, los estimadores funcionales media, mediana y varianza se definen respectivamente como:

$$\bar{X} = \arg \min_{\mathcal{X} \in \mathcal{E}} \sum_{i=1}^n d(x_i, \mathcal{X})^2, \quad (2)$$

$$Med = \arg \min_{\mathcal{X} \in \mathcal{E}} \sum_{i=1}^n d(x_i, \mathcal{X}), \quad (3)$$

$$Var(X) = \sum_{i=1}^n d(x_i, \bar{X})^2. \quad (4)$$

Otro concepto importante en el análisis de datos funcionales es la **profundidad**, la cual mide cuan profundo o central es un punto en relación a un conjunto de puntos. Existen distintas medidas de profundidad: Mediana, Moda, o basada en proyecciones aleatorias. A manera de ejemplo, se presenta la **profundidad modal (PM)** de un conjunto $\{x_1, \dots, x_n\}$, la cual se define como:

$$PM = \max_i g(x_i), i = 1, \dots, n \quad (5)$$

donde:

$$g(x) = \sum_{j=1}^n K \left(\frac{\|x - x_j\|}{h} \right)$$

siendo K una función kernel y h el ancho de banda, el cual se obtiene mediante técnicas de validación cruzada. Existen diferentes tipos de kernel $K(\cdot)$, sin embargo el más usado es el **Gaussiano**: $K(u) = \frac{1}{\sqrt{2}} \exp(-\frac{u^2}{2})$. De manera general,

PM puede ser entendida como una medida de la cantidad de curvas que se encuentran dentro del vecindario de un dato funcional x_i determinado, y por tanto, el punto más profundo será aquel máximo de la función $g(\cdot)$ sobre la muestra. En contraposición, un dato funcional se puede considerar atípico si este minimiza un criterio de profundidad seleccionado.

Posteriormente, y dependiendo de la finalidad del estudio, se pueden utilizar distintas técnicas como la regresión funcional, los modelos de clasificación, entre otras. Por ejemplo, supongamos que se tiene una muestra $(\mathcal{X}, Y) \in \mathcal{E} \times \mathcal{G}$, tal que $\mathcal{G} = \{1, \dots, G\}$, donde cada dato funcional x_i pertenece exclusivamente a un grupo g_i de los G posibles grupos. Los **modelos de clasificación supervisada** intentan estimar la probabilidad a posteriori de pertenencia a los grupos:

$$p_g(x) = P(Y = g | \mathcal{X} = x) = \mathbb{E}(1_{Y=g} | \mathcal{X} = x).$$

La regla de clasificación asignará una nueva observación al grupo que maximice la probabilidad anterior:

$$\hat{Y}(x) = \arg \max_{g \in \mathcal{G}} \hat{p}_g(x)$$

Luego, los modelos de clasificación suelen recurrir a métodos de regresión funcional con variables respuestas binarias, como las regresiones NP (tipo kernel), KNN (k- vecinos cercanos), y GLM (Generalized Linear Model), para construir las reglas de clasificación.

En el presente documento se pretende brindar un panorama general de algunas de las técnicas presentadas anteriormente, restringidas al espacio de las funciones \mathcal{L}_2 . Los resultados son obtenidos a través del software estadístico *R* [12] y sus paquetes, tales como el paquete *fd.usc* [6].

II-A. Ratios financieros de las Cooperativas de Ahorro y Crédito del Ecuador

Las Cooperativas de Ahorro y Crédito (COAC) del Ecuador, se han consolidado como un segmento importante dentro del sistema financiero, en especial a partir de 1999-2000, período en el cual el país se vio sometido a fuertes devaluaciones, congelamiento de depósitos y el proceso de dolarización de la economía. Esta coyuntura influyó fuertemente en los clientes, quienes desplazaron sus movimientos financieros de los bancos hacia las COAC. Por tal razón, ya desde 1984, se establecieron ciertas normativas bajo las cuales las COAC que cumplan ciertos requisitos (en especial relacionados al monto de sus activos) pasaban bajo control directo de la Superintendencia de Bancos y Seguros (SBS). Esto implicaba ser sometidas a rigurosos análisis financieros por parte de auditores externos, con mecanismos de control similares a los aplicados en la banca tradicional [11]. Uno de los resultados de este control, era otorgar una calificación o score de riesgo, la cual establece una clasificación de cada COAC en función de sus niveles operativos y características contables y financieras.

De la información reportada por las COAC a la SBS, esta última construye distintos indicadores y variables para realizar sus reportes y análisis. Entre estas se incluyen el estudio de dos índices financieros: *morosidad* (medido como el porcentaje del saldo de créditos en mora respecto a la cartera total de créditos [19]; y, el *margen neto de utilidad* (representado como

Cuadro I
CUADRO DE AGRUPACIÓN POR CALIFICACIÓN DE RIESGO

Grupo	Calificación	Características de agrupación
1	AA-, A+, A, A-	No presenta problemas significativos
2	BBB+, BBB, BBB-	Presenta problemas menores
3	B+, B, B-, BB+, BB-	Tiene varios problemas significativos

porcentaje de ganancia o pérdida respecto al total de ingresos). En este trabajo, se analiza desde el punto de vista del análisis funcional el comportamiento de estos ratios (que denominaremos *morosidad* y *utilidad* respectivamente), con el objetivo de construir un modelo de clasificación supervisada para cada índice, y luego comparar sus resultados para determinar cuál de ellos asocia de mejor manera la calificación de riesgo a cada COAC.

Con este fin, se han tomado los datos mensuales de estos dos ratios para todas las 38 cooperativas que se encontraban bajo el control de la SBS durante el período de julio 2011 a diciembre 2012, y la calificación de riesgo asociada a cada una de ellas al final de dicho período [18]. Debido al número limitado de cooperativas, se ha procedido a agrupar estas calificaciones de riesgo en tres grupos, de acuerdo con el número (o el nivel de la probabilidad de ocurrencia) de problemas significativos en la gestión de las COAC detectadas por los analistas. Este agrupamiento se presenta en la tabla I.

El esquema de trabajo a seguir en este documento se puede resumir en los siguientes pasos:

1. Realizar la representación funcional de las variables de estudio (y sus derivadas).
2. Cálculo de estadísticos funcionales globales y por grupo de riesgo.
3. Construcción de un modelo de clasificación supervisada, en base a la agrupación por clasificación de riesgo.

III. ANÁLISIS EXPLORATORIO DE DATOS FUNCIONALES

III-A. Representación funcional

Desde el punto de vista funcional, se disponen de dos variables funcionales: *morosidad* y *utilidad*, y en cada una de las cuales se tienen $n = 38$ curvas (correspondientes a cada COAC) con $d = 18$ puntos de discretización equidistantes (uno por cada registro mensual). Por otra parte, las curvas de morosidad tienen un rango definido en $T = [0, 1]$ (pues se trata de un porcentaje, tal como se reporte en los boletines de [19]), mientras que el rango de los datos funcionales de la segunda variable puede tomar tanto valores positivos como negativos, pues el margen neto de utilidad puede corresponder tanto a la ganancia como a la pérdida neta del ejercicio mensual. El conjunto total de variables en la Figura 1, en el cual las curvas de color negro corresponde al grupo 1 ($n_1 = 15$ curvas), el rojo al grupo 2 ($n_2 = 10$ curvas) y el verde al grupo 3 ($n_3 = 13$ curvas).

En esta figura se puede verificar que para la morosidad, las curvas con posiciones muy altas pertenecen en su mayoría al grupo 3 (color verde) y a medida que la curva se encuentra a menor altura, su probabilidad de pertenecer a un grupo con menor nivel de riesgo aumenta. Este comportamiento no ocurre respecto a las curvas de utilidad, en el cual se observa

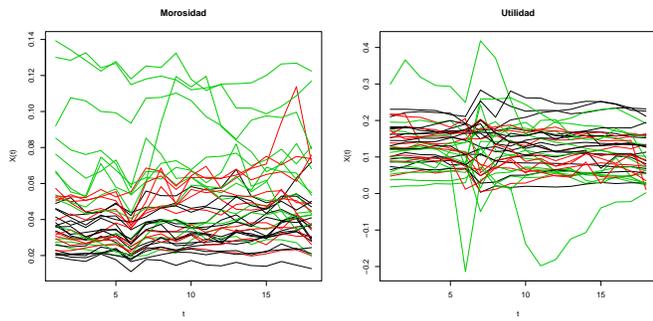


Figura 1. Representación funcional de las curvas muestrales de las variables morosidad y utilidad

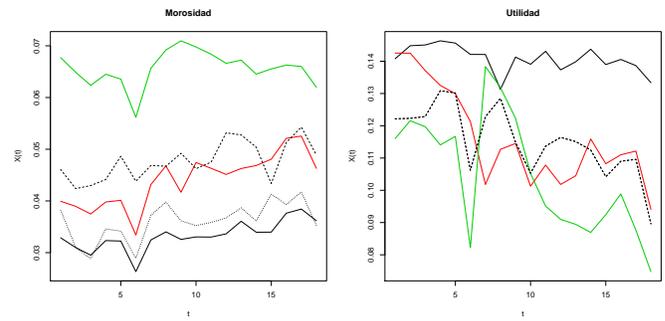


Figura 3. Curvas medias para los grupos 1, 2 y 3 (en líneas continuas negras, roja y verde respectivamente), y curvas (COAC) media y mediana global (en líneas negra discontinua y de puntos respectivamente) para las variables morosidad y utilidad

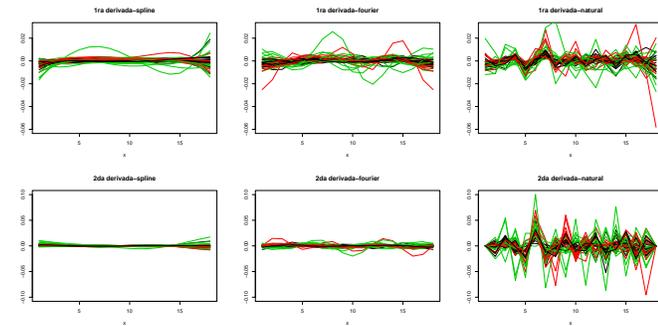


Figura 2. Primera y segunda derivada de la morosidad funcional, representada en bases Bspline, Fourier y Natural

Cuadro II
CURVA (COAC) MEDIA POR GRUPO DE RIESGO PARA MOROSIDAD Y UTILIDAD

Grupo	Morosidad	Utilidad
1	Alianza del Valle	CACPE Pastaza
2	Codesarrollo	San José
3	9 de Octubre	Cotocollao

que todos los grupos se distribuyen al azar dentro del rango observado.

El siguiente paso es seleccionar el espacio funcional adecuado sobre el cual se representarán los datos. En el presente estudio, se ha considerado trabajar en el espacio funcional de \mathcal{L}_2 , pues el área entre curvas puede dar indicios que permitan su posterior clasificación de riesgo. También es necesario verificar si los datos requieren de alguna transformación funcional que permita discriminar mejor los grupos. Para esto se comparan la primera y segunda derivada de la representación en bases de los datos (dada por Eq. 1). Se consideraron tres bases funcionales: Bspline, Fourier y Natural. En la figura 2 se muestran los resultados para morosidad, en la cual se verifica que estas bases no permiten una mejor representación de los diferentes grupos, a diferencia de los datos originales. Este comportamiento también se presentó para la variable utilidad.

III-B. Análisis exploratorio funcional

Para analizar los principales descriptivos, se han tomado los datos originales de cada variable y no su representación en bases, para evitar que los efectos de estas aproximaciones afecten los resultados. En primer lugar, se calculan las curvas medias (dadas por Eq. 2) de cada grupo para cada variable funcional, usando utilizando la métrica establecida para el espacio funcional \mathcal{L}_2 . Además, utilizando los criterios de minimización basados en dicha distancia, es factible determinar que curva (COAC) se aproxima más a la media y a la mediana global (usando Eq. 3) de cada variable. Los resultados obtenidos se observan en la Figura 3.

Para el caso de la morosidad, se verifica que las medias de cada grupo presentan una tendencia ascendente, es decir que conforme aumente el nivel de riesgo, la media funcional de la morosidad toma valores más altos. Asimismo, es importante indicar que la curva media global corresponde a la COAC Andalucía (línea negra discontinua), mientras que la mediana global es de la COAC Oscus (línea negra de puntos), ambas pertenecientes al grupo 1.

Por otra parte, observando el comportamiento de las curvas medias de utilidad de cada grupo se puede verificar que conforme sube el riesgo, las medias tienden a tomar valores más bajos, aunque este comportamiento es más errático en comparación con la morosidad, en especial respecto a la media del grupo 3 (línea verde). Otro aspecto interesante, es que para esta variable, las curvas medias y mediana globales coinciden, y corresponden a la COAC Riobamba (línea negra discontinua), ubicada también en el primer grupo.

Este análisis se aplicó también para obtener las curvas media y mediana dentro de cada grupo. Las curvas (COAC) medias resultantes se presentan en la Tabla II. Respecto a las medianas funcionales, estas coinciden con la curva media para todos los agrupamientos, con excepción del grupo 3 de la variable utilidad, en la cual la curva mediana corresponde a la COAC Santa Ana.

El siguiente paso fue la detección de curvas (COAC) atípicas, para lo cual es necesario utilizar una medida de profundidad. En nuestro caso, utilizamos la profundidad modal definida en la Eq. 5. De manera intuitiva, una curva con PM mínima se puede considerar atípica, pues es la que tiene la menor cantidad de datos dentro de sus vecindarios. Revisando los resultados que se presentan en la figura 4, se identificó que la curva atípica tanto para la morosidad como para la utilidad corresponde a la misma cooperativa, COAC COOPAD.

El siguiente paso en el análisis exploratorio funcional es

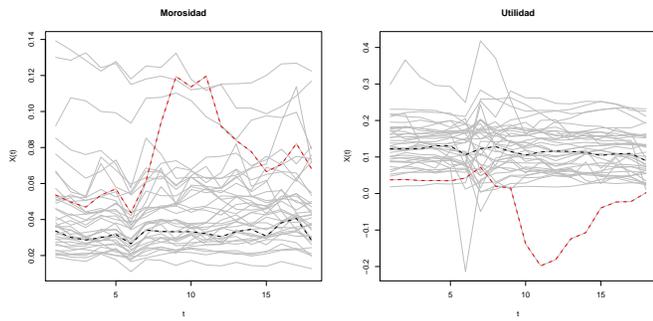


Figura 4. Curva atípica (en línea roja discontinua) y curva mediana global (en línea negra discontinua) para las variables morosidad y utilidad

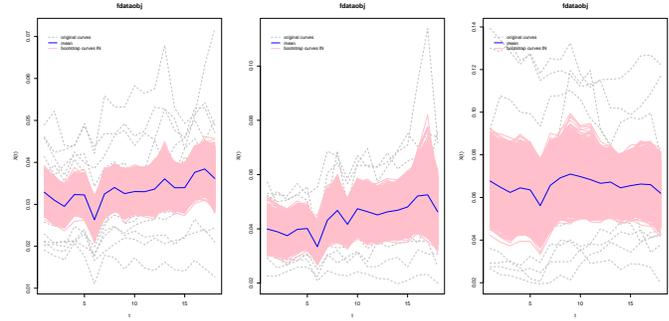


Figura 6. Bandas de confianza al 95 % para la media funcional de morosidad por grupo de riesgo.

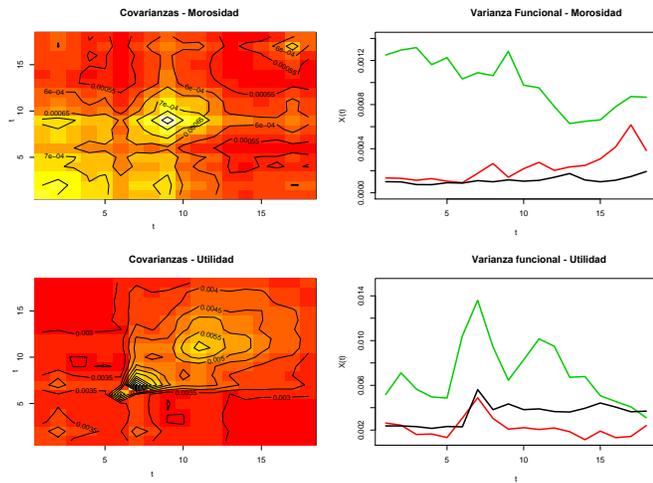


Figura 5. Matriz de Covarianza temporal (izquierda) y varianzas por grupo (derecha) para morosidad y utilidad.

verificar el comportamiento de la dispersión de los datos, en base a la Eq. 4., los cuales se muestran en la figura 5. Para esto, se calcularon la matriz de covarianza temporal (izquierda) y la varianza funcional de cada grupo (derecha) tanto para la morosidad como para la utilidad (primera y segunda fila respectivamente). Respecto a la covarianza de la morosidad, se verifica que la morosidad tiene menor variabilidad entre primeros los datos temporales (meses) y luego esta aumenta para los últimos meses. Esto no sucede con la utilidad, donde se verifica que la covariabilidad es alta en la mayoría de casos. Respecto a las varianzas funcionales por grupo, en ambas variables se observa que la variabilidad del grupo 3 es mayor frente a los otros grupos, aunque este comportamiento es más evidente para la morosidad.

Finalmente, para analizar la variabilidad de los estimadores funcionales calculados, se suele recurrir a la construcción de bandas de confianza Bootstrap. A manera de ejemplo, en la figura 6 se presentan las bandas de confianza al 95 % para la media funcional de cada uno de los tres grupos de la variable morosidad.

Cuadro III
TABLAS DE CONFUSIÓN DE LOS MODELOS DE CLASIFICACIÓN SUPERVISADA NP, KNN Y GLM

Método	Grupos	Morosidad			Utilidad		
		1	2	3	1	2	3
NP	1	14	7	4	13	9	13
	2	1	2	2	0	0	0
	3	0	1	7	2	1	0
KNN	1	15	5	1	15	4	3
	2	0	5	1	0	6	2
	3	0	0	11	0	0	8
GLM	1	11	4	2	11	4	2
	2	2	4	1	1	5	2
	3	2	2	10	3	1	9

IV. MODELOS DE CLASIFICACIÓN SUPERVISADA FUNCIONAL

Los modelos de clasificación supervisada permiten identificar grupos de características homogéneas en base a la información de una muestra, en la cual estos grupos se encuentran plenamente identificados. En nuestro ejemplo, los grupos corresponden a los presentados en la Tabla I.

Desde la perspectiva funcional, para construir los modelos de clasificación supervisada, se tomaron varios métodos de clasificación: NP (tipo kernel), KNN (k- vecinos cercanos), y GLM (Generalized Linear Model). Cada método se aplicó con la variable funcional morosidad y luego con la variable utilidad, utilizando la métrica anterior. Para comparar los resultados de cada modelo, se construyó la tabla de confusión III, donde se muestran la frecuencia de conteo f_{ij} de clasificación (correcta o erróneamente) al momento de asignar a un grupo fila i una curva que pertenece verdaderamente al grupo columna j .

De acuerdo a los resultados obtenidos, se puede establecer que el mejor método de clasificación es el método KNN y el menos eficiente es el NP. Analizando los resultados por grupos, se puede verificar que el grupo 1 es el que presenta menos errores de asignación, mientras que las curvas que corresponden al grupo 2 tienen mayor probabilidad de ser asignadas al grupo 1. En general, el grupo 3 es el que se clasifica incorrectamente con mayor frecuencia. Asimismo, se observa que la clasificación resulta mejor utilizando la variable morosidad con todos los modelos utilizados, lo cual refleja un mejor poder predictivo de dicho ratio frente a la utilidad.

Todos estos resultados son consistentes con los obtenidos

anteriormente en el análisis exploratorio, en el cual el grupo 3 presentaba un comportamiento más disperso en comparación con el resto de grupos, en especial con la variable utilidad.

V. CONCLUSIÓN

Luego de realizar los procesos de análisis funcional, se puede concluir que la variable que mejor permite clasificar las COAC de acuerdo a los niveles de riesgo, es la morosidad en contraste con el margen de utilidad. Esto debido no solamente al comportamiento funcional de esta variable, sino también a que esta variable tiene mayor relación con las diferentes medidas de riesgo financiero que el margen de utilidad.

Al respecto, se evidenció que a mayor valor del ratio de morosidad de una COAC, le correspondía un grupo con mayor nivel de riesgo. Por otra parte, los cálculos relacionados a las curvas medias, profundidad y dispersión indicaban en casi todos los casos que los grupos presentaban mejores características discriminantes con dicho índice que con la variable utilidad. Igualmente, se observó que de manera general, el comportamiento de los grupos con menor riesgo es muy similar entre sí, mientras que el grupo con mayor nivel de riesgo también presenta un comportamiento más disperso, lo que influyó en los resultados de los modelos de clasificación.

Si bien estas conclusiones son usuales en el análisis financiero, lo relevante de estos resultados se basan en que fueron obtenidos mediante el uso de herramientas de FDA en este contexto, lo cual da una medida de su posible congruencia con los resultados proporcionados por otras técnicas con fines similares, además de ser consistentes con los enfoques teóricos de la gestión de riesgos financieros.

Asimismo, y considerando que el alcance del presente trabajo consiste en una aproximación de algunas de las técnicas básicas del análisis funcional, a futuro se puede plantear extender y aplicar otras herramientas, como la regresión funcional, que permitiría obtener mayores y más interesantes conclusiones sobre este conjunto de datos. Por ejemplo, para mejorar el nivel predictivo de los modelos de clasificación se podría levantar más datos de las COAC en distintos períodos, para verificar si es factible realizar predicciones apropiadas, tomando el conjunto de datos actual como muestra de entrenamiento. Sin embargo, no fue factible conseguir más datos, debido a que luego del período de análisis, las cooperativas dejaron de pertenecer a la SBS debido a cambios en la normativa legal.

Finalmente, y a la vista de los resultados obtenidos, se espera que el presente trabajo motive a los distintos académicos, analistas, instituciones financieras y organismos de control, para aplicar estos métodos estadísticos, de manera complementaria con las herramientas existentes en el ámbito financiero.

AGRADECIMIENTOS

El primer y el tercer autor desean agradecer a la Universidad de las Fuerzas Armadas ESPE, por el apoyo brindado para el presente trabajo. De igual manera, el segundo autor agradece a la Escuela Politécnica Nacional por su valioso aporte en la realización de este trabajo.

REFERENCIAS

- [1] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankrupt. *The Journal of Finance*, 589-609.
- [2] Cardot, H.; Faivre, R.; Goulard, M., and Vialard (2001). Functional approaches for predicting land use with the temporal evolution of coarse resolution remote sensing data. *Journal of Applied Statistics*, 1185-1199.
- [3] Jón Danielsson (2002); The emperor has no clothes: Limits to risk modelling. *Journal of Banking & Finance*, Vol. 26, Issue 7, 1273-1296
- [4] De Lara Haro, A. (2005). Medición y control de riesgos financieros. Editorial Limusa.
- [5] Elizondo, A., and Altman, E.I. (2003). Medición integral del riesgo de crédito. Editorial Limusa.
- [6] Febrero-Bande, M., and Oviedo de la Fuente, M. (2012). Statistical Computing in Functional Data Analysis: The R Package fda.usc. *Journal of Statistical Software*.
- [7] Ferraty F, Vieu P. (2006). *Nonparametric Functional Data Analysis*. Springer-Verlag, New York.
- [8] García, M. L. S., and García, M. J. S. (2010). Modelos para medir el riesgo de crédito de la banca. *Cuadernos de Administración*, 295-319.
- [9] Hippler, W., and Hassan, K. (2015). The impact of macroeconomic and financial stress on the U.S financial sector. *Journal of Financial Stability*, 61-80.
- [10] Lizaraburu, E., Berggrun, L., and Quispe, J. (2012). Gestión de riesgos financieros. Experiencia en un banco latinoamericano. *Estudios Gerenciales*, 96-103.
- [11] Miño Grijalva, W. (2013). *Historia del Cooperativismo en Ecuador*. Quito: Editogran S.A.
- [12] R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, (2008). <http://www.R-project.org>.
- [13] Ramsay, J. O., and Dalzell, C. J. (1991). Some Tools for Functional Data Analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 539-572.
- [14] Ramsay, J. O., and Silverman, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer-Verlag.
- [15] Ramsay, J., and Silverman, B. (2006). *Functional Data Analysis*. Springer-Verlag.
- [16] Ratcliffe, S. J., Leader, L., and Heller, G. Z. (2002). Functional data analysis with application to periodically stimulated foetal heart rate data. I: Functional regression. *Statistics in Medicine*, 1103-1114.
- [17] Ravi Kumar, P., and Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques A review. *European Journal of Operational Research*, 128.
- [18] Superintendencia de Bancos del Ecuador. (2013). Calificación de Riesgo de Instituciones Financieras. Recuperado en Diciembre de 2013, de http://www.sbs.gob.ec/practg/sbs_index?vp_art_id=836&vp_tip=2
- [19] Superintendencia de Bancos del Ecuador. (2013). Boletines Financieros mensuales de las Cooperativas de ahorro y crédito del Ecuador, de http://www.superbancos.gob.ec/practg/sbs_index?vp_art_id=488&vp_tip=2&vp_busr=41#2
- [20] Silva, W., Kimura, H., and Amorin-Sobreiro, V. (2016). An Analysis of the literature on systemic financial risk: a survey. *Journal of Financial Stability*, DOI:10.1016/j.jfs.2016.12.004.
- [21] Tam, K. (1991). Neural Network Models and the Prediction of Bank Bankruptcy. *Omega*, 429-445.

Sergio Castillo-Páez, is a mathematical engineer graduated from the National Polytechnic School in 2002. He also studied finance in the Simon Bolivar Andean University, and is currently studying his PhD in Statistics at the University of Vigo, Spain. He is a professor at ESPE - University of the Armed Forces in Ecuador.



His current lines of research are related to geostatistics and analysis of multivariate data.



Miguel Flores is a professor at the National Polytechnic School and a researcher at the Center for Modeling Mathematics at the National Polytechnic School in Quito, Ecuador. He is a BSc. in Statistical Computing Engineer from the Polytechnic School of the Coast. In 2006 he received an MSc. in Operations Research from the National Polytechnic School, and in 2013 received a MSc. in Technical Statistics from the University of A Corua. He is currently a doctoral student at the University of A Corua in the area of Statistics and Operations Research. He has over 15 years professional experience in various areas of Statistics, Computing and Optimization, multivariate data analysis, econometric, Market Research, Quality Control, definition and construction of systems indicators, development of applications and optimization modeling. ORCID ID: 0000-0002-7742-1247



Giovanni Herrera-Enrriquez is a professor at ESPE - University of the Armed Forces in Ecuador. He is a Ph.D in Economics and Business from the University of Santiago de Compostela(Spain) and he has a Master degree in economics and business management by FLACSO. His current lines of research are related to models for decision making, natural disaster risk management, multicriteria hierarchical analysis and market research.