

## TÉCNICAS DE CALIFICACIÓN DE PRÉSTAMOS: LA METODOLOGÍA ROUGH SET COMO ALTERNATIVA AL ANÁLISIS DISCRIMINANTE

M.<sup>a</sup> DOLORES OLIVER ALFONSO\*

REYES SAMANIEGO MEDINA\*\*

M.<sup>a</sup> JOSÉ VÁZQUEZ CUETO\*\*\*

### INTRODUCCIÓN

En 1999, El Comité de Basilea, consciente de los cambios que se han ido produciendo en el mundo bancario en los últimos años, se planteó una reestructuración del Acuerdo de 1988 sobre la medición y control de los riesgos de las entidades financieras. Así, en junio de 2004, se aprueba el documento definitivo del nuevo Acuerdo: Basilea II. Acuerdo que marcará un punto de partida tanto en la gestión de los riesgos como en las relaciones que habrán de mantener entidades financieras y supervisores.

Basilea II obliga así, a la implantación de modelos para la medición de los riesgos financieros en las entidades de crédito. Estos modelos tienen como objetivo el cálculo de la pérdida inesperada (IL) de las exposiciones crediticias. Esta pérdida depende de un conjunto de factores: probabilidad de impago (PD), pérdida en caso de impago (LGD), exposición en el momento del impago o severidad (EAD), vencimiento (M) y granularidad (G).

Nuestro trabajo se estructura en tres partes. En la primera presentamos los aspectos más sobresalientes del tratamiento del riesgo de crédito en Basilea II. En la segunda parte realizamos una revisión de la literatura financiera. Y por último presentamos la metodología de los rough sets, que «enfrentamos» al análisis discriminante clásico en una aplicación empírica en la que tratamos de determinar cuál o cuáles son las variables que pueden actuar como explicativas de la morosidad.

\* Universidad de Sevilla.

\*\* Universidad Pablo de Olavide.

\*\*\* Universidad de Sevilla.

## EL RIESGO DE CRÉDITO EN EL NUEVO ACUERDO DE BASILEA

El tratamiento de los riesgos financieros se ha convertido en un factor estratégico, ya no sólo para las entidades financieras, sino para cualquier organización, con independencia del tamaño que posea y del sector en que realice su actividad. Si nos centramos en las entidades financieras, el incremento de la competencia, así como los avances producidos en la diversificación y en la liquidez y los cambios tan significativos en la regulación, tales como los requerimientos de capital basados en el riesgo, han llevado a las entidades a buscar formas innovadoras que ayuden a la medición y gestión de todos los riesgos que pueden influir en su desarrollo. Así se ha recuperado el interés por un riesgo, que aún existiendo desde los comienzos de la actividad bancaria, no se le había prestado la atención que le correspondía. Estamos hablando del riesgo de crédito.

El riesgo de crédito ha sido el gran olvidado por la banca. La falta de interés por este tipo de riesgo refleja el hecho de que las técnicas para su medición y control no hayan evolucionado prácticamente nada en el último siglo, al contrario de lo ocurrido con el riesgo de mercado, el cual ha experimentado un importante auge tanto en su estudio como en su análisis, originando que las entidades financieras cuenten con modelos matemáticos para su gestión. No obstante, la realidad actual es bien distinta, tanto académicos como profesionales han empezado a dar al riesgo de crédito la importancia que se merece, a la vez que es frecuente encontrar referencias a éste en la prensa especializada. A pesar de que las entidades financieras han asumido la gran importancia que el riesgo posee, su estudio y medición no resultan fáciles, por lo que el camino que queda aún por recorrer es largo. Así como al medir el riesgo de mercado podemos hacer estudios de las variables que intervienen en él —al existir datos históricos para ello— cuando hablamos del riesgo de crédito, la cuestión se complica debido a la escasez de información histórica a analizar.

El objetivo básico del nuevo Acuerdo es proporcionar un cálculo del capital bancario necesario más sensible al riesgo y para ello propone la utilización de metodologías internas de medición del riesgo elaborada por los propios bancos. La determinación del capital mínimo regulador supone una modificación importante del Acuerdo de 1988. Establece métodos para el cálculo del capital regulador que cubran el riesgo de crédito. Se preconizan dos métodos, el *estándar*, que es una mejora del utilizado en el acuerdo anterior, y que establece nuevas categorías de riesgo, agrupando a cada tipo de empresa dentro de éstas y el Método Interno o IRB (Internal Rating Based) basado en calificaciones internas. Para utilizar el método IRB es necesario poder calcular la probabilidad de incumplimiento (PD). Para ello se tiene que calcular la PD media de todos los prestatarios incluidos en un grupo o grado de clasificación. Estas se obtendrán como resultado de las estimaciones, basadas en datos históricos, de las funciones que modelizan los comportamientos, mostrando siempre una visión conservadora a medio y largo plazo.

## MODELIZACIÓN DEL RIESGO DE CRÉDITO EN LA LITERATURA FINANCIERA

Los modelos tradicionalmente utilizados son los denominados «Modelos de forma reducida»<sup>1</sup>, en los que se considera la insolvencia como un fenómeno exógeno. Todo el esfuerzo teórico se centra en la modelización del proceso estocástico asociado a la misma y en la determinación de las variables que habrán de ser introducidas. En este sentido podemos distinguir entre los modelos univariantes, que estudian el comportamiento de cada una de las variables potencialmente explicativas de la insolvencia por separado<sup>2</sup>, y los modelos multivariantes, que combinan la información proporcionada por un conjunto de variables, siendo el trabajo pionero en este sentido, el realizado por Altman en 1968<sup>3</sup>. Más recientemente, Cabedo, Reverte y Tirado (2004) hacen una adaptación de este modelo para el cálculo de las probabilidades de impago en corporaciones, sobre una cartera hipotética perteneciente a un mismo sector, con el objeto de calcular el capital regulatorio según el método IRB básico del Acuerdo de Basilea. En las últimas décadas, sin embargo, se ha optado fundamentalmente por el análisis logit<sup>4</sup>. Así, Wilson (1997) desarrolló el modelo CreditPortfolio View para Mckinsey, obteniendo las probabilidades de impago como funciones logísticas de índices de variables macroeconómicas que, de alguna manera, representan el funcionamiento de la economía. Más recientemente, Fernandez (2005) utiliza la misma metodología realizando previamente un análisis univariante con en fin de elegir, de entre los 23 ratios de partida, aquellos con mayor poder discriminante, dentro de cada una de las categorías de ratio establecidas<sup>5</sup>. Posteriormente obtiene puntuaciones para cada individuo que le permiten establecer un sistema de rating y asignar probabilidades de impago. Junto con esto, en la última década del pasado siglo, y en un intento de superar las hipótesis tan restrictivas que imponen los modelos de inferencia estadística a las variables, hipótesis que no se suelen cumplir en la realidad y que distorsionan los resultados de las técnicas<sup>6</sup>, se empiezan a aplicar técnicas procedentes del campo de la Inteligencia Artificial, elaborando programas capaces de generar conocimiento a través de datos empíricos y, posteriormente, usar ese conocimiento para realizar inferencias sobre nuevos datos. Dentro de este enfoque podemos distinguir las técnicas que buscan el conocimiento a través de anticipar patrones en los datos, entre ellas las diversas redes neuronales, y las consistentes en inferir reglas de decisión a partir de los datos de la base. A este último grupo pertenece la metodología de los Rough Sets<sup>7</sup>.

1. Desarrollado, entre otros, por Jarrow y Turnbull (1995) y Duffie y Singleton (1999).

2. Uno de los estudios clásicos a este respecto fue realizado por Beaver (1966), el cual encontró que un número de ratios financieros podría discriminar entre empresas sanas y fallidas, en un periodo de 5 años antes de que se produjera el impago.

3. Altman (1968) realizó un análisis discriminante combinando la información proporcionada por 25 ratios financieros.

4. El primero en usar este tipo de técnicas aplicada a la predicción empresarial fue Ohlson (1980).

5. Liquidez, apalancamiento, actividad, cobertura de deuda y productividad.

6. Eiseimbeis (1977), Ohlson (1980) y Zavgren (1983) cuestionan por estas razones la validez del método.

7. Daubie et al (2002) han aplicado dicha técnica a la clasificación de préstamos comerciales.

## METODOLOGÍA ROUGH SETS

La teoría Rough Set fue propuesta por Z. Pawlak en 1982 y se ha afianzado como una buena herramienta para tratar con la vaguedad y la imprecisión. Más concretamente, el método es eficiente para tratar con los problemas de clasificación multidimensionales (Pawlak y Slowinski, 1994). La idea básica descansa en la relación de indiscernibilidad que describe elementos indistinguibles. Su objetivo básico es encontrar reglas de decisión básicas que permitan adquirir nuevo conocimiento. Sus conceptos básicos son 1) la discernibilidad, 2) la aproximación, 3) los reductos y 4) las reglas de decisión. El punto de partida del método es la existencia de una tabla de información/decisión, donde cada elemento viene caracterizado por un conjunto de variables y una variable de decisión, que clasifica al elemento en una de dos ó más categorías. La indiscernibilidad se produce cuando dos elementos vienen caracterizados por los mismos valores de las variables y, sin embargo, no coinciden las categorías en las que se clasifican. Esta es la base de los conjuntos rough set. En tal caso, y para cada clase de decisión o categoría  $X$ , y cada subconjunto  $B$  de variables, se construyen dos conjuntos, denominados respectivamente aproximaciones por debajo y aproximaciones por encima de la clase de decisión. El conjunto aproximación por debajo de la clase de decisión  $X$  respecto a las variables  $B$ ,  $\underline{B}X$  viene dado por el conjunto de todos los elementos que, caracterizados por  $B$ , con toda seguridad pertenecen a  $X$ . El conjunto aproximación por encima de la clase  $X$ ,  $\overline{B}X$  viene dado por el conjunto de elementos que, en base a la información  $B$  que poseemos, no podemos asegurar que no estén en  $X$ . La diferencia entre los dos conjuntos la forman los elementos «dudosos», es decir, aquellos que no sabemos con toda certeza, usando sólo la información contenida en  $B$ , si pertenecen o no pertenecen a  $X$ . Cuando esta diferencia es no vacía se dice que  $X$  es un rough set respecto a  $B$ . Este conjunto puede ser caracterizado numéricamente por el cociente entre el cardinal del conjunto aproximación por debajo y el cardinal del conjunto aproximación por encima. A este cociente se le denomina precisión. Si existen varias clases de decisión, a la suma de los cardinales de todas las aproximaciones por debajo dividida por el total de elementos se denomina calidad de la clasificación por medio del conjunto  $B$ . Es el porcentaje de elementos correctamente clasificados.

Otro aspecto importante es la reducción de la tabla intentando eliminar aquella información redundante. Este proceso se realiza a través de los reductos. Un reducto es un conjunto mínimo de variables que conserva la misma habilidad de clasificación de los elementos que la tabla completa de información. Un reducto es pues la parte esencial del conocimiento y constituyen la manera más concisa mediante la cual podemos distinguir las clases de decisión<sup>8</sup>.

La etapa final del análisis es la creación de reglas de decisión, es decir, reglas que nos permitan decir si un elemento dado pertenece a determinadas clases de decisión. Estas reglas representan el conocimiento y se generan combinando los

8. La obtención de los reductos está basada en las clases de equivalencia que define, sobre el conjunto de las observaciones, la relación de indiscernibilidad.

reductos con los valores de los datos. Una regla de decisión es una sentencia lógica del tipo «SI(se cumplen determinadas condiciones) ENTONCES (el elemento pertenece a una determinada clase de decisión)». Estas reglas nos permitirán, de manera fácil, clasificar a nuevos elementos<sup>9</sup>.

## APLICACIÓN EMPÍRICA

Como el objetivo de este trabajo es mostrar la utilidad de este método, alternativo al análisis discriminante, en la etapa previa a la calificación de un préstamo, determinando las variables que mejor discriminan a los prestatarios, hemos realizado una aplicación empírica. Para llevarla a cabo hemos empleado una base de datos suministrada por una caja de ahorros española, que contiene información de empresas que solicitaron y obtuvieron un crédito en dicha entidad. Estas compañías son divididas en dos grupos: sanas y fallidas. En concreto, la muestra de empresas fallidas utilizada para el análisis ha sido obtenida de aquellas compañías, prestatarias de la entidad financiera, cuya deuda resultó impagada, de intereses o principal, en un porcentaje superior al 10% del riesgo concedido. El momento de cómputo ha sido el 31 de diciembre de 2003.

El grupo de empresas sanas, es decir aquellas que no generaron situaciones de impago en el horizonte considerado, se ha seleccionado mediante la técnica del emparejamiento individual, controlado por aquellas características que podrían incidir en las relaciones entre ratios financieros y fracaso. A cada empresa del grupo de fallidas, se ha unido una empresa sana de la misma industria y de aproximado tamaño. En relación al sector, el emparejamiento se ha realizado a nivel de cuatro dígitos del C. N. A. E. de 1993. El criterio seguido para emparejar por tamaño ha sido el activo total.

Como factor homogeneizador a todas las empresas, hemos controlado que el total de operaciones del cliente con la entidad financiera, o riesgo vivo, fuese superior a 60.120 euros, y que se tratase de sociedades anónimas o limitadas, lo que facilitaría el acceso a sus estados contables.

En total, la muestra se compone de 106 empresas, 53 fallidas y 53 sanas de las más diversas actividades económicas.

Las variables independientes escogidas para la elaboración de los modelos fueron seleccionadas de los estados financieros, fundamentalmente balance y cuenta de pérdidas y ganancias, de las empresas que componen la muestra. Dichos estados contables fueron extraídos de la base de datos SABE (Informa, Información Económica, S. A.) que incluye más del 95% de las compañías que presentan sus cuentas en el Registro Mercantil en España. Dado que la mayoría de las empresas que resultaron fallidas no presentaron los estados financieros el año anterior o incluso en los dos años previos a la fecha de impago, hemos considerado los últimos datos disponibles como los del año previo al fallo empresarial. Así, el año t-1 corresponde al último de los estados contables disponibles.

9. Para más detalle sobre los aspectos matemáticos formales de la metodología puede consultarse Komorowski et al (1999)

La selección de ratios se ha hecho escogiendo un conjunto amplio de variables, 25, potencialmente explicativas del fracaso empresarial sobre la base de la frecuencia y eficacia con que han sido utilizadas en otros modelos de predicción de la insolvencia empresarial o en el análisis de riesgos bancarios, siguiendo la clasificación realizada por Trujillo (2002).

Las variables utilizadas incluyen ratios de liquidez, endeudamiento, estructura, rotación, generación de recursos y rentabilidad. En concreto, los ratios considerados en el análisis se recogen en el cuadro 1.

Cuadro 1. *Ratios considerados en el análisis*

Ratios de Liquidez	Grado en que los activos liquidables a corto plazo de la empresa son suficientes para hacer frente a los pagos provenientes de las deudas contraídas a corto plazo.	R1=Activo Circulante/ Pasivo Circulante R2=(Exigible+Disponible)/Pasivo Circulante R3=Disponible/Pasivo Circulante R4=(Exigible+Disponible-Pasivo Circulante)/(Consumos de Explotación+Gastos de Personal + Variación Provisiones + Otros Gastos de Explotación)
Ratios de Endeudamiento	Relación entre los distintos componentes del pasivo, a corto y largo plazo y los recursos propios; así como entre el coste de la deuda y el pasivo o los beneficios y recursos generados	R5=Pasivo Fijo/Neto R6=Neto/Pasivo Total R7=Pasivo Fijo/(Pasivo Fijo + Pasivo Circulante) R8=Costes Financieros/(Pasivo Fijo + Pasivo Circulante) R9=Costes Financieros/(BAIT + Dotación Amortización) R10=Costes Financieros/BAIT R11=Pasivo Ajeno/Pasivo Total
Ratios de Estructura	Proporcionalidad entre las masas patrimoniales del activo y del pasivo, así como en la composición de las mismas	R12=(Activo Circulante-Pasivo Circulante)/Activo Total. R13=Activo Circulante/Activo Total
Ratio de Rotación	Medida del dinamismo de la actividad empresarial en relación a la estructura de la empresa.	R14=(Activo Circulante-Pasivo Circulante)/(Importe Neto Cifra de Negocio + Otros Ingresos de Explotación)
Ratios de Generación de Recursos	Relación de la capacidad de autofinanciación de la empresa con diferentes magnitudes.	R15=(Resultado Ejercicio + Dotación Amortización) / (Importe Neto Cifra Negocios + Otros Ingresos de Explotación) R16=(Resultado Ejercicio + Dotación Amortización)/ Pasivo Circulante R17=(Resultado Ejercicio + Dotación Amortización)/ (Pasivo Fijo + Pasivo Circulante) R18=(Resultado Ejercicio + Dotación Amortización)/ Pasivo Total R19=(BAIT+Dotación Amortización)/Pasivo Circulante
Ratios de Rentabilidad	Comparación del beneficio obtenido en sus distintos niveles, con los recursos invertidos	R20=(Resultado Explotación + Ingresos Financieros + Beneficios Inversiones Financieras + Diferencias Positivas de Cambio)/Activo Total R21=Resultado de Actividades Ordinarias/Pasivo Total R22=Resultado antes de Impuestos/Neto R23= Resultado antes de Impuestos/Pasivo Total R24= Resultado del Ejercicio /Neto R25=BAIT/Activo Total

Se han calculado los valores de 25 ratios económico – financiero (cuadro 1)<sup>10</sup> para cada una de las 53 empresas fallidas, un periodo antes de la entrada en mora, y se ha procedido de igual forma para cada «pareja» sana. Se obtiene así una tabla de información de 106x25 datos. A esta tabla se le ha unido una columna indicativa de la situación de fracasada o saneada de la empresa en cuestión. Así hemos asignado el valor 0 a la empresa fallida y 1 a la correspondiente pareja sana. Obtendremos, de esta forma, una tabla de información-decisión de 106x26 datos.

Con esta tabla tratamos de determinar cuál o cuáles de estos 25 ratios pueden actuar como explicativos de la morosidad, y así, actuar de «sistemas de alerta» preventiva para las entidades financieras.

A continuación, y dada la naturaleza de las variables consideradas, hemos procedido a la discretización de los valores. Esto no es requisito imprescindible para la aplicación de la técnica, pero nos ha parecido, al igual que a otros autores, que facilitaba la interpretación de los resultados y que parecía más coherente identificar países con sistemas bancarios en crisis o saneados, no cuando coincidieran exactamente los valores de las variables consideradas, sino cuando éstos se encontraran dentro de un mismo intervalo. Para ello hemos utilizado la codificación que se recoge en el cuadro 2<sup>11</sup>.

Cuadro 2. Intervalos de codificación de las variables

Var./valor cod.	0	1	2	3
R1	(-inf, -0.226568)	(-0.226568, -0.217696)	(-0.217696, -0.14786)	(-0.14786, +inf)
R2	(-inf, -0.594852)	(-0.594852, -0.359847)	(-0.359847, 0.854273)	(0.854273, +inf)
R3	(-inf, -0.342112)	(-0.342112, -0.336467)	(-0.336467, 0.032813)	(0.032813, +inf)
R4	(-inf, 0.0930971)	(0.0930971, 0.319831)	(0.319831, 0.336117)	(0.336117, +inf)
R5	(-inf, -0.132725)	(-0.132725, -0.109552)	(-0.109552, -0.087447)	(-0.087447, +inf)
R6	(-inf, -0.408355)	(-0.408355, 0.0850476)	(0.0850476, 0.120947)	(0.120947, +inf)
R7	(-inf, -1.1135)	(-1.1135, -1.05743)	(-1.05743, -0.85543)	(-0.85543, +inf)
R8	(-inf, -0.652671)	(-0.652671, -0.599027)	(-0.599027, 0.0167526)	(0.0167526, +inf)
R9	(-inf, -0.331797)	(-0.331797, -0.189594)	(-0.189594, 0.059217)	(0.059217, +inf)
R10	(-inf, -0.350028)	(-0.350028, -0.100267)	(-0.100267, 0.127363)	(0.127363, +inf)
R11	(-inf, -0.348336)	(-0.348336, -0.0322035)	(-0.0322035, 0.48369)	(0.48369, +inf)
R12	(-inf, -1.01652)	(-1.01652, 0.106868)	(0.106868, 0.193528)	(0.193528, +inf)
R13	(-inf, -1.22338)	(-1.22338, -0.069475)	(-0.069475, 0.025187)	(0.025187, +inf)
R14	(-inf, -0.152482)	(-0.152482, 0.174396)	(0.174396, 0.597593)	(0.597593, +inf)
R15	(-inf, -0.207683)	(-0.207683, -0.115229)	(-0.115229, 0.763632)	(0.763632, +inf)
R16	(-inf, -0.464693)	(-0.464693, -0.301845)	(-0.301845, -0.255046)	(-0.255046, +inf)
R17	(-inf, -0.511584)	(-0.511584, -0.311845)	(-0.311845, 0.326507)	(0.326507, +inf)
R18	(-inf, -0.210892)	(-0.210892, -0.193191)	(-0.193191, -0.003212)	(-0.003212, +inf)
R19	(-inf, -0.768045)	(-0.768045, -0.767186)	(-0.767186, -0.453772)	(-0.453772, +inf)
R20	(-inf, -0.2324)	(-0.2324, 0.0165318)	(0.0165318, 1.75011)	(1.75011, +inf)
R21	(-inf, -0.261301)	(-0.261301, 0.0323638)	(0.0323638, 1.78434)	(1.78434, +inf)
R22	(-inf, -0.183217)	(-0.183217, -0.163807)	(-0.163807, 0.237888)	(0.237888, +inf)
R23	(-inf, -0.06379)	(-0.06379, -0.04726)	(-0.04726, 0.103724)	(0.103724, +inf)
R24	(-inf, -0.162962)	(-0.162962, -0.1209)	(-0.1209, 0.256906)	(0.256906, +inf)
R25	(-inf, -0.491449)	(-0.491449, -0.465705)	(-0.465705, -0.20096)	(-0.20096, +inf)

Elaboración propia.

10. Se ha trabajado con los logaritmos neperianos de los mismos.

11. Para ello se ha utilizado el software ROSE proporcionado por el Institute of Computing Science of Poznan University of Technology, al que agradecemos su disponibilidad.

El siguiente paso es la construcción de los reductos. Al existir correlaciones entre las variables explicativas introducidas en el análisis, el número de reductos que ofrecía el programa era muy elevado. Optamos por establecer como criterio de parada los veinticinco primeros, así aparecieron los reductos que se recogen en el cuadro 3, cada uno de los cuales contiene entre siete y nueve variables, lo que ya supone una reducción con respecto a las veinticinco originales. La frecuencia con la que aparece cada variable en los reductos viene dada en el cuadro 4.

Cuadro 3. *Reductos*

---

1: {R8, R11, R13, R15, R16, R18, R20, R24, R25}
2: {R10, R11, R13, R16, R22, R24, R25}
3: {R10, R11, R13, R17, R19, R22, R24, R25}
4: {R10, R11, R13, R18, R19, R22, R24, R25}
5: {R10, R11, R13, R19, R20, R22, R24, R25}
6: {R10, R11, R13, R14, R19, R24, R25}
7: {R10, R11, R13, R19, R23, R24, R25}
8: {R7, R9, R15, R16, R17, R22, R23, R24, R25}
9: {R9, R13, R14, R15, R16, R21, R22, R24}
10: {R9, R11, R14, R15, R16, R20, R22, R24}
11: {R9, R14, R15, R16, R18, R19, R20, R22, R24}
12: {R9, R14, R15, R16, R18, R19, R21, R22, R24}
13: {R9, R11, R14, R15, R16, R21, R22, R24}
14: {R9, R14, R15, R16, R18, R20, R21, R22, R24}
15: {R9, R13, R14, R15, R16, R19, R22, R24}
16: {R9, R14, R15, R16, R19, R21, R22, R23, R24}
17: {R9, R14, R15, R16, R18, R19, R22, R23, R24}
18: {R9, R11, R14, R15, R16, R21, R22, R23, R25}
19: {R9, R14, R15, R16, R17, R20, R22, R23, R25}
20: {R11, R12, R13, R14, R16, R22, R23, R25}
21: {R10, R11, R13, R19, R21, R24, R25}
22: {R11, R12, R13, R15, R16, R22, R24, R25}
23: {R10, R11, R13, R17, R21, R22, R23, R24}
24: {R9, R14, R15, R16, R20, R22, R23, R24}
25: {R11, R13, R14, R15, R16, R22, R24, R25}

---

Cuadro 4. *Aparición de las variables en los reductos*

---

<i>Variable</i>	<i>Frecuencia</i>	<i>Variable</i>	<i>Frecuencia</i>
R1	0%	R7	4%
R2	0%	R8	4%
R3	0%	R9	52%
R4	0%	R10	32%
R5	0%	R11	60%
R6	0%	R12	8%

---

Cuadro 4. Aparición de las variables en los reductos (Cont.)

Variable	Frecuencia	Variable	Frecuencia
R13	56%	R20	28%
R14	60%	R21	32%
R15	64%	R22	84%
R16	72%	R23	36%
R17	16%	R24	88%
R18	24%	R25	56%
R19	44%		

*Elaboración propia.*

Esto nos indica que las variables más discriminatorias en la muestra son R24 y R22, ambos ratios de rentabilidad, que aparecen en más del 80% de los reductos seleccionados, lo que significa que no podemos prescindir de ellas sin renunciar a parte de información de la situación. Les sigue la variable R16 que aparece en los reductos con una frecuencia superior al 70%. Llama la atención el hecho de que los seis primeros ratios no han sido seleccionados para formar parte de ninguno de los reductos. Recordemos que los cuatro primeros se corresponden con los ratios de liquidez. Podemos, pues, prescindir de ellos sin perder información, deducimos, así, que la mayor o menor liquidez de las empresas no es determinante de la morosidad de la misma. De entre todos los reductos que ha proporcionado el software hemos seleccionado el 6, {R10, R11, R13, R14, R19, R24, R25}. Los criterios seguidos para la selección han sido: 1º que contuviese el menor número de variables posibles, 2º que las variables presentasen una frecuencia elevada de aparición en los reductos, 3º que contuviese ratios económicos y financieros<sup>12</sup>

Con estas variables hemos reducido la tabla de información/decisión a una tabla 106x8 datos y hemos obtenido las veintiocho reglas de decisión que figuran en el cuadro 5.

Cuadro 5. Reglas de decisión.

regla 1. Si (R10 = 3) y (R13 = 3) y (R19 = 2) entonces (D1 = 0);
regla 2. Si (R10 = 3) y (R14 = 2) entonces (D1 = 0);
regla 3. Si (R10 = 3) y (R13 = 3) y (R14 = 1) y (R24 = 2) entonces (D1 = 0);
regla 4. Si (R13 = 0) y (R14 = 0) entonces (D1 = 0);
regla 5. Si (R14 = 3) y (R25 = 2) entonces (D1 = 0);
regla 6. Si (R11 = 2) y (R13 = 3) y (R14 = 1) y (R19 = 2) y (R25 = 3) entonces (D1 = 0);
regla 7. Si (R10 = 2) y (R14 = 3) entonces (D1 = 0);
regla 8. Si (R13 = 0) y (R14 = 1) y (R24 = 1) entonces (D1 = 0);
regla 9. Si (R11 = 2) y (R14 = 0) y (R25 = 2) entonces (D1 = 0)
regla 10. Si (R13 = 1) y (R19 = 2) y (R25 = 2) entonces (D1 = 0);
regla 11. Si (R24 = 3) entonces (D1 = 0);

12. Estas condiciones las cumplen también los reductos 2, 7 y 21, que sustituyen el ratio R15 por ratios de rentabilidad (R22, R23 y R21 respectivamente). Nos ha parecido más conveniente mantener R15 porque es el único ratio de rotación que intervendrá en la clasificación, mientras que otros ratios de rentabilidad, como lo son R24 y R25, ya están introducidos.

Cuadro 5. Reglas de decisión (Cont.)

---

regla 12. Si $(R11 = 0)$ y $(R14 = 0)$ entonces $(D1 = 0)$ ;
regla 13. Si $(R24 = 0)$ entonces $(D1 = 0)$ ;
regla 14. Si $(R10 = 1)$ y $(R13 = 3)$ y $(R14 = 1)$ y $(R19 = 3)$ entonces $(D1 = 0)$ ;
regla 15. Si $(R10 = 1)$ y $(R14 = 2)$ entonces $(D1 = 1)$ ;
regla 16. Si $(R10 = 1)$ y $(R13 = 1)$ entonces $(D1 = 1)$ ;
regla 17. Si $(R10 = 2)$ y $(R14 = 1)$ y $(R19 = 3)$ y $(R24 = 2)$ entonces $(D1 = 1)$ ;
regla 18. Si $(R11 = 0)$ y $(R14 = 2)$ entonces $(D1 = 1)$ ;
regla 19. Si $(R13 = 2)$ entonces $(D1 = 1)$ ;
regla 20. Si $(R10 = 2)$ y $(R13 = 3)$ y $(R25 = 2)$ entonces $(D1 = 1)$ ;
regla 21. Si $(R10 = 1)$ y $(R24 = 1)$ y $(R25 = 3)$ entonces $(D1 = 1)$ ;
regla 22. Si $(R14 = 1)$ y $(R24 = 2)$ y $(R25 = 2)$ entonces $(D1 = 1)$ ;
regla 23. Si $(R13 = 1)$ y $(R19 = 2)$ y $(R25 = 3)$ entonces $(D1 = 1)$ ;
regla 24. Si $(R13 = 1)$ y $(R24 = 1)$ y $(R25 = 3)$ entonces $(D1 = 1)$ ;
regla 25. Si $(R14 = 3)$ y $(R24 = 2)$ entonces $(D1 = 1)$ ;
regla 26. Si $(R11 = 1)$ y $(R24 = 1)$ entonces $(D1 = 1)$ ;
regla 27. Si $(R10 = 1)$ y $(R13 = 0)$ y $(R25 = 3)$ entonces $(D1 = 1)$ ;
regla 28. Si $(R13 = 3)$ y $(R14 = 1)$ y $(R19 = 3)$ y $(R24 = 1)$ entonces $(D1 = 1)$ ;

---

*Elaboración propia.*

Las catorce primeras clasifican a las empresas morosas, y las catorce siguientes a las restantes. Así, si el ratio de endeudamiento  $R10 = 3$  y el ratio de rotación  $R14 = 2$ , es decir, y según los criterios de discretización, si el ratio Costes Financieros/BAIT es superior a 1.135 ( $\exp 0.127$ ) y el ratio (Activo Circulante – Pasivo Circulante) / (Importe Neto Cifra de Negocio + Otros Ingresos de Explotación) está comprendido entre los valores 1.1905 y 1.8177 ( $\exp 0.174396$  y  $\exp 0.597593$  respectivamente), la empresa será clasificada como fallida ( $D_1 = 0$ ), (Regla 2). Por el contrario, si el ratio de endeudamiento  $R10 = 1$ , es decir, es mayor que 0.7046 ( $\exp -0.350028$ ), y menor que 0.9046 ( $\exp -0.100267$ ), y el ratio de estructura  $R13 = \text{Activo Circulante} / \text{Activo Total} = 1$ , es decir toma valores entre 0.2942 y 0.93288 ( $\exp -1.22338$  y  $\exp -0.0694753$  respectivamente), entonces la empresa será clasificada como no fallida, ( $D1 = 1$ ), (Regla 16). Estas veintiocho reglas clasifican con un 100% de aciertos a los elementos utilizados. Además si realizamos una validación de los resultados mediante el procedimiento de «dejar uno fuera», seguimos obteniendo un porcentaje aceptable de aciertos, en concreto un 76.42% (un 71.7% para las empresas morosas y un 81.13% para las no morosas).

Para poder comparar este método con el análisis discriminante, utilizamos el mismo conjunto de datos sin discretizar. Construimos dos funciones discriminantes, una utilizando los ratios seleccionados por el modelo usando como criterio de entrada  $F\text{-Snedecor} > 1,5$  y como criterio de salida  $F\text{-Snedecor} < 1$  y otra utilizando sólo las variables del reducto seleccionado con el método rough set. Los coeficientes de estas funciones discriminantes se muestran en la cuadro 6.

El cuadro 7 presenta los resultados de la reclasificación con ambas funciones discriminantes, aplicadas a la muestra original y validándolas mediante el procedimiento de validación cruzada.

Cuadro 6. *Funciones discriminantes y estadísticos de los modelos*

	<i>Lambda de Wilks</i>	<i>P-valor</i>
-0,571 R4-0,687R6+0,312R7+0,559R8+0,274R9+1,074R15-0,354R19-,713R23	0,572	0,0
0,250R10+0,824R11-0,201R13-0,307R14-0,124R19+0,198R24-0,192R25	0,750	0,0

*Elaboración propia*

Cuadro 7. *Porcentajes de clasificación correcta.*

	<i>Análisis Discriminante (ocho variables)</i>		<i>Análisis Discriminante (variables de la metodología rough set)</i>	
Muestra de elaboración	Global	82,1	Global	72,6
	Saneados	84,9	Saneados	75,5
	En crisis	79,2	En crisis	69,8
Muestra de validación	Global	81,1	Global	69,8
	Saneados	84,9	Saneados	75,5
	En crisis	77,4	En crisis	64,2

*Elaboración propia.*

De esta tabla se infiere que la metodología Rough Set se ha mostrado como una herramienta más útil que el Análisis Discriminante. Utilizando menor número de variables, e, incluso, con las mismas variables, no sólo muestra porcentajes globales de clasificación correcta mejores sino que también clasifica mejor dentro de cada una de las clases de decisión. Estos porcentajes disminuyen en la muestra que se ha considerado para la validación de los resultados cuando comparamos la metodología propuesta con los resultados del análisis discriminante clásico, aunque siguen conservándose superiores a los obtenidos con el análisis discriminante cuando restringimos las variables a aquellas que la metodología rough set ha considerado como suficientes para la clasificación.

## CONCLUSIONES

El objetivo básico del Comité de Supervisión Bancaria de Basilea, plasmado en el nuevo Acuerdo, ha sido proporcionar un cálculo del capital bancario necesario más sensible al riesgo. Para ello ha propuesto la utilización de metodologías internas de medición del riesgo elaborada por los propios bancos. Para conseguir este objetivo, sin que se produzca una desregulación, el Comité incluye en el nuevo Acuerdo partes o técnicas hasta ahora no tenidas en cuenta. Así en la propuesta aparecen elementos novedosos tales como la inclusión técnicas internas, con distinto grado de valoración del riesgo de crédito. El Comité ha establecido que las entidades que lo deseen puedan utilizar modelos internos desarrollados para estimar las variables fundamentales de las que depende la valoración del riesgo de crédito.

En nuestro trabajo, que se enmarca en una línea de investigación más amplia encaminada a la determinación de una de las variables fundamentales de los modelos internos: la probabilidad de impago (PD), presentamos una metodología alternativa al clásico análisis discriminante para determinar cuál o cuáles son las variables que pueden actuar como explicativas de la morosidad de una empresa y así, servir de «sistemas de alertas» preventivas para las entidades financieras.

Apoyándonos en una muestra de empresas sanas y fallidas y en un conjunto de 25 ratios financieros potencialmente explicativos de la morosidad de las mismas hemos obtenido que la metodología Rough set puede ser una alternativa válida al análisis discriminante cuando se trata de problemas de clasificación de objetos en dos clases de decisión. Junto con la obtención de porcentajes de clasificación correcta aceptables une el hecho de que no exige de ningún tipo de comportamiento previo de las variables que intervienen en la clasificación, a diferencia del análisis discriminante que exige normalidad de las distribuciones e igualdad de las matrices de varianza-covarianza. Entre sus ventajas podemos destacar también que elimina la información redundante y expresa las dependencias entre las variables consideradas y el resultado de la clasificación mediante unas reglas de decisión cuyo lenguaje está más próximo al lenguaje natural de un experto.

## BIBLIOGRAFÍA

- Altman, E. I. (1968): «Financial Ratios, Discriminant Analysis and the Predictions of Corporate Bankruptcy» *Journal of Finance* (September) 589-609.
- , (1998): *Managing Credit Risk. The New Great Financial Challenge*. John Wiley & Sons, Inc.
- Basle Committee on Banking Supervision (2001): «The Internal Ratings-Based Approach». Consultive Document. Supporting Document to the New Basle Capital Accord, enero.
- Beaver, W. H. (1966): «Financial Ratios as Predictors of Failure». *Empirical Research in Accounting: Selected Studies*, (Suplemento de *The Accounting Review*).
- Carey, M., Hrycay, M. (2001): «Parameterizing Credit Risk Models with Rating Data». *Journal of Banking & Finance*, 25.
- Comité de Supervisión Bancaria de Basilea (2004): «Convergencia Internacional de Medidas y Normas de Capital». junio.
- Cabedo, J., Reverte, J. A., Tirado, J. M. (2004): «Riesgo de Crédito y Recursos Propios Mínmicos en Entidades Financieras». *Revista Europea de Dirección y Economía de la Empresa*, volumen 13, núm. 2.
- Crouhy, M., Galai, D., Mark, R. (2001): «Prototype Risk Rating System». *Journal of Banking & Finance*, 25.
- Daubie, M.; Leveck, P.; Meskens, N.(2002):«A comparison of the rough sets and recursive partitioning induction approaches: An application to commercial loans» *International Transactions in Operational research* 9, 681-694.
- Duffie, D.; Singleton, K.(1999): «Modeling Term Structures of Defaultable Bonds» *Review of Financial Studies*, 4.
- Eisembeis, R. A. (1977):«Pitfalls in the application of discriminant analysis in business and economics» *Journal of Finance*, 32, 875-900.
- Fernández, J. E. (2005): «Corporate Credit Risk Modeling: Quantitative Rating System and Probability of Default Estimation». *DefaultRisk.com*, abril.

- Jarrow, R. A.; Turnbull, S. M.(1995): «Pricing Derivatives on Financial Securities Subject to Credit Risk», *Journal of Finance*, Vol. L, 1, 53-85.
- Komorowski, Z. (1999):»*Rough Sets: A Tutorial, Rough-Fuzzy Hybridization - A New Trend in Decision Making*, (S.K. Pal and A. Skowron, Eds.), 3-98, Springer-Verlag Singapore Pte Ltd.
- Ohlson, J. A. (1980): «Financial ratios and the probabilistic prediction of bankruptcy» *Journal Of Accounting Research* (Spring) 109-131.
- Pawlak Z., Rough sets. *Int. J. Computer and Information Sci.*, 11, 1982, 341-356
- Pawlak, Z.; Slowinski, R. (1994): «Decision analysis using rough sets. *International Transactions in Operational Research* 1, 107-114.
- Trujillo, A. (2002): *Gestión del Riesgo de Crédito en Préstamos Comerciales*. Instituto Superior de Técnicas y Prácticas Bancarias. Madrid.
- Wilson, T. (1997): «Portfolio Credit Risk», *Risk magazine*, Sept., 111-117.
- Zavgren, C. V.(1983): «The prediction of corporate failure: the state of the art» *Journal of Financial Lietrature* 2, 1-37.