



**TESIS DE GRADO EN
INGENIERÍA INDUSTRIAL**

**UN ANÁLISIS ESTADÍSTICO
PARA PREDECIR LA CESACIÓN DE PAGOS**

Autor: Nicolás Emanuel MONTI

Director de Tesis: Ing. Roberto Mariano GARCÍA

2007



DESCRIPTOR BIBLIOGRÁFICO

El objetivo perseguido con este estudio consiste en identificar a las empresas que tienen una alta probabilidad de entrar en cesación de pagos en el corto plazo. A tal fin, se relevaron un total de 45 índices financieros en 86 empresas, la mitad de las cuales se habían presentado en concurso preventivo.

Por un lado, se utilizó el análisis de componentes principales para reducir la información contenida en los 45 índices originales en dos nuevas variables, con las que se puede establecer la evolución que ha tenido una empresa en su operación y su rentabilidad. Por el otro, se aplicó el análisis de regresión logística para desarrollar un modelo que permita estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo. El modelo finalmente elegido logró identificar con éxito al 85% de las empresas de la muestra.



ABSTRACT

The aim of this study is to identify those firms which are likely to become financially distressed in the short term. To accomplish that, a group of 45 financial ratios were obtained from a sample of 86 firms. Half of these firms had already declared themselves as financially distressed.

On the one hand, the principal components analysis was used in order to turn the information in the 45 original ratios into two new variables. In this way, the firm's performance and the return evolution can be established. On the other hand, the logistic regression analysis was applied to develop a model for predicting the likelihood for firms to become financially distressed. The model finally selected managed to successfully identify the 85% of the companies in the sample.



RESUMEN EJECUTIVO

El problema que se intenta resolver en este trabajo se puede formular a través de la siguiente pregunta: ¿es posible determinar si una empresa tiene problemas financieros a partir de la información contenida en sus estados contables?. En primer lugar, debemos definir que significa que una empresa tenga problemas financieros. En nuestro caso, se ha considerado que una empresa alcanza dicho estado cuando tiene una alta probabilidad de entrar en cesación de pagos en el corto plazo.

Para resolver la problemática planteada se ha recurrido al uso de la inferencia estadística. Primero, se utilizó el análisis de componentes principales, y luego, el análisis de regresión logística. El estudio se basa en un conjunto de 45 índices financieros relevados en una muestra compuesta por 86 empresas que operan en la Argentina. La mitad de estas empresas no tuvieron problemas financieros, mientras que el resto ya se había presentado en concurso preventivo.

El análisis de componentes principales permitió reducir la información contenida en los 45 índices originales en dos nuevas variables que se han denominado *Variación Operativa* y *Variación de la Rentabilidad*. Al representar en un mismo gráfico los valores que toman estas componentes en la muestra, se identificaron ciertos sectores que indican como ha evolucionado una empresa en los últimos ejercicios respecto de su operación y su rentabilidad. De esta manera, se puede conocer su predisposición hacia los problemas financieros. Sin embargo, estas variaciones no siempre permiten establecer en forma precisa que tan grave es la situación financiera de una empresa.

Por tal motivo, se utilizó el análisis de regresión logística para estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo. El modelo finalmente seleccionado explica el 65% de la variación total de las empresas de la muestra y está formado por las siguientes variables: (1) *Endeudamiento corriente*; (2) *Costo financiero*; (3) *Margen operativo*; (4) *Variación de la rentabilidad del patrimonio neto*. A través de este modelo, aproximadamente en el 85% de los casos se pudo establecer con éxito si una empresa de la muestra había tenido problemas financieros.

Por lo tanto, los resultados obtenidos indican que la respuesta a la pregunta inicialmente planteada es afirmativa. El uso de estos análisis estadísticos para el procesamiento de la información contable permite identificar a las empresas con problemas financieros. Así se han generado dos herramientas de gran utilidad para el seguimiento de la situación financiera de una empresa y para la toma de decisiones de inversión, entre otras aplicaciones.



EXECUTIVE BRIEF

The problem that we are trying to solve in this paper can be expressed in the following question: Is it possible to deduce if a firm has financial problems from the information in its balance sheet?. In the first place, we need to define what we mean by financial problems. In this case, it has been considered that a firm reaches that state when it is likely to become financially distressed in the short term.

In order to answer this question, statistical inference was used. Firstly, the principal components analysis was applied and then the logistic regression analysis. The study is based on a group of 45 financial ratios which were obtained from a sample of 86 firms operating in Argentina. Half of these firms did not have any financial problems, while the other half had gone through financial distress.

After using the principal component analysis, the information in the 45 original ratios was turned into two new variables, which were called *Operative Variation* and *Return Variation*. When the values of these components were represented in a single graph, different sectors were defined. Each of them indicate how the operation and the return of a firm has evolved in the last years. In this way, a firm's likelihood to get into financial trouble may be determined. However, those variations are not always precise enough to establish how bad the financial situation in a firm is.

Because of that, the logistic regression analysis was used to estimate the likelihood that a firm should become financially distressed in the short term. The model which was finally selected explains the 65% of the total variability of the firms in the sample and it is represented by the following variables: (1) *Current Debt*; (2) *Financial Cost*; (3) *Operating Profit Margin*; (4) *Return on Equity Variation*. Approximately, in 85% of the cases, the financial situation of the firms in the sample was correctly identified using this model.

Therefore, the results obtained indicate that the answer to the initial question is affirmative. Using the statistical analysis to process the information in the financial statements enable us to identify those firms with financial problems. In conclusion, we have developed two new tools. These may be very useful to track down the financial performance of a firm and to make investment decisions, among other applications.



ÍNDICE

1. Introducción	1
2. Diseño de la muestra.....	3
3. Selección de variables.....	9
4. Análisis del perfil de la muestra	11
5. Análisis de Componentes Principales	17
5.1. Introducción.....	17
5.2. Cálculo de las componentes principales	18
5.3. Resultados obtenidos.....	21
6. Análisis de Regresión Logística.....	29
6.1. El modelo de regresión logística	29
6.2. Estimación de los parámetros.....	31
6.2.1. Estimación por mínimos cuadrados	31
6.2.2. Estimación por mínimos cuadrados ponderados	32
6.2.3. Estimación por máxima verosimilitud	32
6.3. Validación del modelo de regresión logística	33
6.3.1. Ensayos de hipótesis	34
6.3.2. Medidas de bondad de ajuste	36
6.3.3. Pruebas de significancia de los coeficientes de regresión.....	38
6.3.3.1. Contraste de Wald.....	38
6.3.3.2. Contraste de Wilks	40
6.4. Resultados Obtenidos.....	41
6.4.1. Primer modelo a partir de las componentes principales	41
6.4.2. Segundo modelo a partir de las variables originales	48
7. Pruebas con una nueva muestra de datos	61
8. Conclusiones.....	65
9. Bibliografía	67
10. Anexos	69
10.1. Composición detallada de la muestra de empresas	69
10.2. Descripción de los índices financieros considerados.....	70
10.2.1. Índices de liquidez.....	70
10.2.2. Índices operativos	71
10.2.3. Índices de endeudamiento	73
10.2.4. Índices de cobertura.....	74
10.2.5. Índices de rentabilidad	74
10.2.6. Índices de evolución.....	76
10.2.7. Otros índices	78
10.3. Rentabilidades del Patrimonio Neto por tipo de actividad.....	79
10.4. Autovalores de las componentes principales.....	81
10.5. Valores de las componentes principales retenidas.....	81
10.6. Correlaciones entre las componentes principales y las variables.....	83
10.7. Modelos de regresión logística evaluados	83
10.8. Composición detallada de la nueva muestra de empresas	85





1. INTRODUCCIÓN

El objetivo que se persigue con este estudio consiste en identificar a aquellas empresas que presentan problemas financieros a partir de la información contenida en sus estados contables. A tal efecto, se considera que una empresa se encuentra en este estado cuando tiene una alta probabilidad de entrar en cesación de pagos en el corto plazo.

El análisis se basa en el uso de la inferencia estadística sobre un conjunto de índices financieros relevados en empresas que se han presentado en concurso preventivo¹ y en otras que no lo han hecho.

En algunos trabajos similares se propuso la búsqueda de un índice que tuviera la mayor capacidad para identificar a una empresa con problemas financieros, o bien, de un modelo estadístico que mejor pueda predecir la quiebra de una empresa mediante el análisis discriminante. En nuestro caso, se aplicará el análisis de componentes principales y el de regresión logística con el fin de desarrollar un conjunto de herramientas que puedan utilizarse de manera complementaria para la toma de decisiones en situaciones de inversión, otorgamiento de créditos y gerenciamiento, entre otras.

En la sección 2 se lleva a cabo el diseño de la muestra, donde se definen su tamaño y composición así como los criterios utilizados para el relevamiento de la información. La muestra está compuesta por 86 empresas que se dividen en dos grandes grupos, aquellas que no han tenido problemas financieros (grupo 1) y aquellas que se han presentado en concurso preventivo (grupo 2).

En la sección 3 se establece el conjunto de variables a utilizar que constituye la base de los análisis posteriores. En total se han relevado 45 índices financieros que representan diferentes características de las empresas, relacionadas con la liquidez, la operación, el endeudamiento, la cobertura, la rentabilidad y la evolución de otros índices.

En la sección 4 se realiza un análisis del perfil de la muestra que consiste en graficar los histogramas de ciertos índices representativos con el fin de evaluar su comportamiento en los dos grupos de empresas considerados.

En la sección 5 se efectúa el análisis de componentes principales a través del cual se pretende resumir la información contenida en los 45 índices financieros en sólo unas pocas nuevas variables.

¹ El concurso preventivo es un instrumento legal que presupone la cesación de pagos en las empresas.



En la sección 6 se desarrollan diferentes modelos de regresión logística con el objetivo de estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo.

En la sección 7 se aplican las herramientas basadas en el análisis de componentes principales y de regresión logística sobre una nueva muestra de datos con el objetivo de evaluar la efectividad conjunta en el reconocimiento de las empresas con problemas financieros.

Por último, en la sección 8 se muestran las conclusiones del estudio así como sus posibles aplicaciones y algunas alternativas para continuar con esta investigación.



2. DISEÑO DE LA MUESTRA

Un aspecto muy importante en este tipo de trabajos es la selección y recolección de la información que constituye la muestra de datos a partir de la cual se va a construir el modelo estadístico. Por ejemplo, si se trabaja con una muestra de empresas que pertenecen únicamente a la industria de la construcción, es de esperar que el modelo que se obtenga sólo sea aplicable a empresas de dicho sector. O bien, si la muestra está constituida por un 90% de empresas que no han tenido problemas financieros y sólo un 10% de empresas que se declararon en cesación de pagos, el poder discriminatorio del modelo resultante no será significativo. Por tal motivo, a continuación se comentan los criterios utilizados en el diseño de la muestra que determinarán el alcance del análisis.

La muestra está constituida por un total de 86 empresas que operan en la Argentina de las cuales 43 no han presentado problemas financieros (grupo 1) mientras que otras 43 se han presentado en concurso preventivo en el período analizado (grupo 2). En el Anexo 10.1 se muestra una composición detallada de la muestra.

La información utilizada en el presente estudio se relevó a partir de los estados contables. En el caso de las empresas no concursadas, estos se obtuvieron de la Bolsa de Comercio de Buenos Aires (BCBA), mientras que para el caso de las empresas concursadas se consultaron los informes generales de los síndicos que intervinieron en los concursos respectivos².

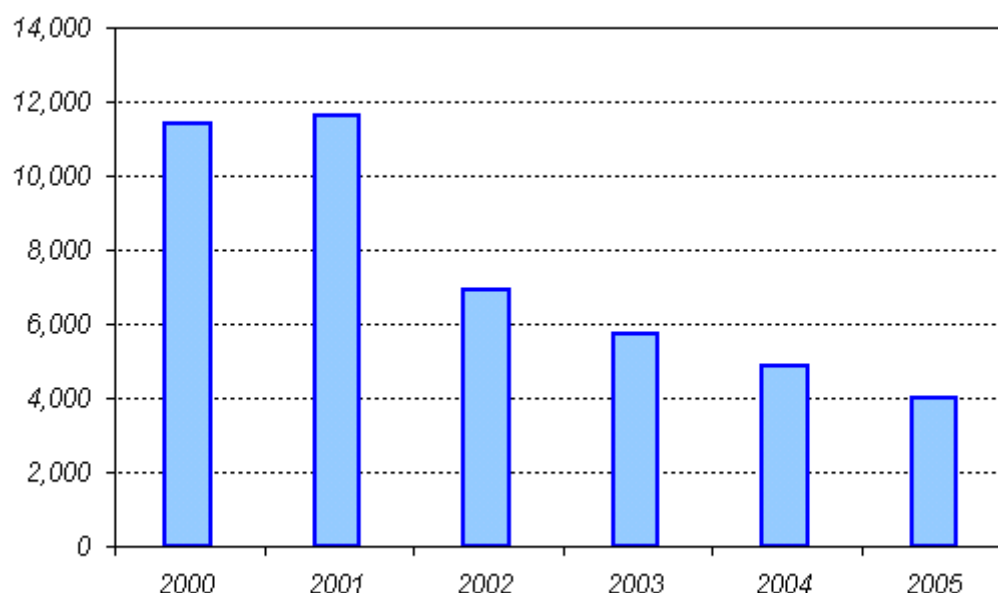
Diversos autores de libros de estadística consideran válido relevar un mínimo de 5 observaciones por cada variable explicativa que se desee incluir en el modelo. En los trabajos de William Beaver [1] y Edward Altman [2], se adoptaron muestras de mayor y de menor tamaño - 120 y 60 empresas respectivamente - y se obtuvieron buenos resultados, habiendo considerado en ambos casos no más de 5 variables explicativas. Esto indica, teniendo en cuenta que en este análisis no se espera trabajar con modelos que tengan más de 5 variables explicativas, que una muestra de 86 empresas es razonable.

En cuanto a la proporción de empresas con y sin problemas financieros, si bien no es estrictamente necesario tomar la misma cantidad de observaciones para cada grupo, es recomendable respetar esta condición para lograr una mejor representación de los valores medios y la dispersión de las variables observadas en cada grupo. Para entender fácilmente esta cuestión, podemos analizar el caso extremo de una muestra compuesta por 1 empresa que no ha presentado problemas financieros y 99 empresas que sí lo han hecho. Llegado el momento de calcular la probabilidad de que una empresa entre en cesación de pagos, es razonable pensar que el modelo tendrá una marcada tendencia a clasificar a dicha empresa como si fuera a tener problemas

² Los informes son publicados mediante el sistema informático de la Cámara Nacional de Apelaciones en lo Comercial.



financieros. Esto se debe a que la muestra, al no ser representativa de la población, no “revela” las distintas maneras en que puede presentarse una empresa financieramente sana. Es decir, la muestra contiene muy poca información del comportamiento de las variables observadas en empresas sin dificultades financieras, y por lo tanto, es más difícil que el modelo reconozca a una empresa de dicho grupo.



Fuente: Cámara Nacional de Apelaciones en lo Comercial

Figura 2.1 – Evolución de los expedientes concursales ingresados en Cap. Fed.

Otra cuestión importante para analizar se relaciona con los períodos de tiempo a partir de los cuales se elige recopilar la información, sobre todo en el caso de las empresas concursadas. La muestra utilizada incluye empresas que han operado durante los años 2003, 2004 y 2005 en ambos grupos. Es importante destacar que si el período a partir del cual se toma la muestra es demasiado amplio, por ejemplo mayor a 10 años, se corre el riesgo de mezclar información contable de empresas que operan en un contexto macroeconómico distinto. Esto último sugiere una interpretación particular de los índices financieros de cada empresa, inclusive de un mismo sector. En países que presentan una estabilidad económica duradera este efecto no introduciría una gran distorsión en la muestra de datos, pero éste no es el caso de la Argentina. Al respecto, se tomó la decisión de no incluir en el estudio la información contable de empresas que se hayan presentado en concurso durante los años 2001 y 2002, debido a que en dicha época aconteció una crisis económica en el país que afectó el normal desenvolvimiento de las empresas. De esta manera, se procura no introducir en la muestra una variación atípica, que no es objeto de estudio en este análisis y que podría distorsionar los resultados esperados. Sin embargo, hay que tener en cuenta que en el caso de las empresas que se presentaron en concurso durante el 2003, sí se tomó en cuenta la información contable del 2002 con el objetivo de contar con una



evolución de los índices financieros. De todas formas, como se puede observar en la Figura 2.1, la evolución de la cantidad total de expedientes concursales ingresados en los tribunales comerciales de la Capital Federal (donde se encuentra el domicilio legal de la mayor parte de las empresas que operan en nuestro país) muestra una marcada caída a partir del 2002. Esto permite suponer que la variación atípica causada por la crisis que se introduce en la muestra al incorporar información contable del período 2002-2003 no es significativa.

El procedimiento adoptado para la recopilación de la información, en el caso de las empresas con problemas financieros, consistió en relevar los estados contables correspondientes al ejercicio económico en que esas empresas se presentaban en concurso preventivo y del año anterior, de manera de incluir en el análisis la evolución de los índices financieros entre dos ejercicios (ver sección 3. Selección de variables). En el caso de las empresas que no se presentaron en concurso, simplemente se escogieron los balances correspondientes a dos ejercicios consecutivos dentro del período bajo estudio.

Respecto de este tema, cabe advertir que en estudios similares se incluyó en el análisis la información contable de hasta 5 años previos a la declaración de la quiebra. Sin embargo, en general, en dichos estudios se utilizó la información contable de cada año por separado, en vez de considerar variables que reflejen la evolución de los estados contables en una única muestra. La metodología de análisis empleada consistió en utilizar la información contable de años anteriores como muestras separadas para testear la capacidad de discriminación de un modelo obtenido a partir de los índices financieros recopilados en el período más próximo a la declaración de la quiebra. Los resultados obtenidos reflejaron, como era de esperar, que a medida que la información contable de una empresa se alejaba más en el tiempo respecto del momento en que se produjo la quiebra, la capacidad del modelo de distinguir entre empresas con o sin problemas financieros disminuía. Esto demuestra que no resulta relevante incluir en el análisis información contable de varios años previos a la presentación en concurso de las empresas, dado que en ese entonces los índices financieros pueden reflejar un buen desempeño que dificulta su clasificación. Al contrario, resulta razonable prestar especial atención a la actividad de la empresa en el momento en el cual las características distintivas se hacen evidentes, es decir, en los últimos años previos a la presentación en concurso.

Las empresas incluidas en la muestra para este trabajo pertenecen a diferentes sectores de la economía que incluyen la industria, el comercio, el agro y la prestación de servicios. El principal motivo de dicha elección consiste en desarrollar un modelo abarcativo que pueda ser aplicado en distintos tipos de empresas.

La teoría financiera establece que no pueden compararse en forma directa los índices financieros de dos empresas que pertenecen a sectores distintos de la economía, dado que la dinámica de cada uno de ellos no es la misma. Por ejemplo, una empresa financieramente sana puede mostrar un índice de liquidez absoluta de 2 en un



determinado sector, mientras que otra empresa que desarrolla otro tipo de actividad económica puede tener el mismo valor de dicho índice y encontrarse en graves problemas financieros. Por lo tanto, a primera vista no parecería razonable incluir empresas de diferentes sectores, porque la muestra podría contener contradicciones sobre las características que permiten identificar a una empresa con problemas financieros, es decir, se estaría distorsionando la relación entre los índices financieros y la cesación de pagos. Sin embargo, debemos recordar que al tratarse de un análisis multivariante, las características que se observan en cada individuo se comparan en forma global y simultánea, por lo tanto, es más difícil que los comportamientos particulares de algunos índices en determinados sectores de la economía puedan distorsionar el perfil global de una empresa que la califica dentro de alguno de los dos grupos considerados. De todas formas, se pueden tomar ciertas precauciones con el fin de disminuir la incidencia del sesgo propio de cada sector económico. La primera de ellas consiste en incluir en ambos grupos los mismos sectores, evitando así que haya un grupo con empresas de un sector que no esté presente en el otro. Al mismo tiempo, la situación mejoraría si la cantidad de empresas de un sector en cada grupo es la misma. Aunque esta premisa no se haya logrado a la perfección, debido a la dificultad que representa la recopilación de la información, se buscó dentro de lo posible un equilibrio en la muestra.

William Beaver, en su trabajo *Financial Ratios as Predictors of Failure* [1], preparó una muestra apareada sobre la base de empresas que operaban en diferentes sectores de la industria. En dicha muestra, por cada empresa que había mostrado incumplimiento en sus pagos se incluía otra financieramente sana de la misma industria y en lo posible del mismo tamaño. En ese entonces, Beaver desarrolló un análisis estadístico univariante, es decir, que se compararon los índices de cada empresa uno a la vez y la discriminación entre empresas financieramente sanas de aquellas que no lo eran se basaba en un único índice a través de un punto de corte.

En su investigación, este autor propuso efectuar un *análisis apareado* con el objetivo de cuantificar si el efecto que produce el sesgo de cada sector industrial y el tamaño de las empresas en los índices es importante. A tal fin, consideró para cada par de empresas del mismo sector y de tamaño similar la diferencia entre los índices financieros relevados y luego evaluó si esas diferencias resultaban lo suficientemente grandes como para ubicar a las empresas en alguno de los dos grupos considerados. Dado que todas las diferencias entre índices se calculan comparando dos empresas del mismo sector y de tamaño similar, las posibles distorsiones que estos factores pueden introducir en la muestra se estarían mitigando.

Cabe destacar que este análisis sólo fue realizado a los efectos de cuantificar la influencia que ejercen sobre los índices el comportamiento particular de cada sector de la economía y el tamaño de las empresas. Sin embargo, para encontrar el punto de corte en cada índice que permite la posterior clasificación de las empresas, se compararon directamente los valores observados de los índices. Esto se debe, fundamentalmente, al hecho de que con un análisis apareado no se pueden sacar



conclusiones sobre un individuo en particular, es decir, siempre se deben comparar dos empresas en forma simultánea.

Al contrastar los resultados obtenidos en el análisis apareado y en la comparación global de las empresas, se esperaría que los errores de clasificación calculados para cada índice mediante ambos análisis sean similares, siempre y cuando el efecto residual propio de cada industria y del tamaño de las empresas en la muestra seleccionada sean despreciables. Las evidencias encontradas avalan esta última hipótesis. Por lo tanto, todo parece indicar que la utilización de una muestra apareada es la mejor opción para mitigar estos efectos. Sin embargo, debemos tener en cuenta que dicho estudio se basó en un análisis estadístico univariante que no permitía comparar varios índices al mismo tiempo. Por esta razón, no es estrictamente necesario contar con una muestra apareada para continuar con nuestro estudio, aunque buscar cierto equilibrio contribuye a disminuir la distorsión que provocan en los índices las particularidades de cada sector de la economía y el tamaño de las empresas.

Otra medida interesante que se ha implementado en este trabajo con el fin de facilitar el reconocimiento de empresas con problemas financieros en diferentes sectores de la economía, radica en la inclusión de una variable que mide la performance de una empresa respecto de otras que desarrollan la misma actividad. En la próxima sección se muestra con más detalle la totalidad de las variables consideradas.

Por último, un aspecto importante que hay que tener en cuenta en relación con el diseño de la muestra - que se ha mencionado ya al comentar el estudio de Beaver - es el tamaño de las empresas observadas, considerando a éste último como la totalidad de activos que dichas empresas poseen. En nuestro caso, se ha tomado la decisión de no incluir en la muestra a empresas de gran tamaño. Esto se debe, fundamentalmente, al hecho de que son muy pocos los casos en los cuales las grandes empresas se han presentado en concurso preventivo, y por lo tanto, es factible pensar que pertenecen a una población distinta a la de las empresas con problemas financieros.

Alexander Sydney, en su trabajo *The Effect of Size of Manufacturing Corporation on the Distribution of the Rate of Return* [3], comenta que existen fundamentos teóricos así como evidencias empíricas que demuestran que la tasa de retorno de una empresa se torna más estable a medida que aumenta su tamaño. Esto podría implicar que una empresa de gran tamaño tendría un riesgo menor de entrar en cesación de pagos respecto de una empresa mediana o pequeña, inclusive si ambas exhiben los mismos valores en los índices financieros. Por lo tanto, se podría pensar, en principio, que no se pueden comparar directamente los índices financieros de dos empresas cuyos tamaños son muy distintos. Teniendo en cuenta que para llevar a cabo un análisis consistente es necesario que las empresas de ambos grupos provengan de la misma población, se ha procurado que el rango de variación del tamaño de las empresas en ambos grupos sea similar. De todas maneras, buscar una homogeneidad perfecta en la muestra en lo que refiere al tamaño de las empresas, no es deseable,



porque estaríamos reduciendo la capacidad de discriminación y predicción del modelo en su conjunto.



3. SELECCIÓN DE VARIABLES

En la literatura sobre la materia se puede encontrar una gran cantidad de índices financieros que se definen en función de los distintos aspectos que se desean conocer de una empresa. La selección de las variables que luego van a ser sometidas al análisis estadístico comprende una etapa muy importante de la investigación, dado que en este momento se deben contemplar todas aquellas características de las empresas que uno piensa que pueden incidir causalmente en la cesación de pagos. En cierta forma, al igual que en el diseño de la muestra, aquí se define el alcance y la validez del estudio. En nuestro caso, dicha selección se llevó a cabo según los siguientes criterios: (1) la popularidad que poseen los índices financieros en la literatura; (2) el desempeño que han demostrado ciertos índices en estudios anteriores; (3) una valoración propia del autor en cuanto a la utilidad de un determinado índice como discriminador de empresas con problemas financieros.

Los análisis que se llevarán a cabo consideran un total de 45 variables, cuyos valores se han calculado para cada una de las empresas incluidas en la muestra según los criterios detallados en la sección anterior. La mayoría de estas variables son índices financieros, aunque algunas de ellas corresponden a indicadores de crecimiento, o bien, muestran la evolución entre dos ejercicios económicos de otros índices.

En el Anexo 10.2 se puede encontrar una breve descripción del significado de cada índice así como también una mención de aquellas variables populares en la literatura que fueron originalmente consideradas para formar parte del análisis pero que luego, por diferentes motivos, se decidió descartarlas. A pesar de que se han excluido todas aquellas relaciones que sean una transformación de otro índice, muchos de ellos poseen numeradores y denominadores en común pero cuya interpretación es distinta.

Asimismo, se ha incluido un nuevo índice denominado *Rentabilidad Comparada* con el objetivo de establecer una comparación entre la rentabilidad del patrimonio neto de cada empresa y la rentabilidad promedio de otras empresas del mercado que desarrollan la misma actividad. En el Anexo 10.3 se muestra un detalle de las rentabilidades promedio utilizadas para el cálculo de este índice.

Inicialmente se ha considerado un gran número de variables explicativas de manera de contar con una base de datos completa que nos permita desarrollar y comparar diferentes modelos de regresión. A su vez, teniendo en cuenta que se llevará a cabo un análisis de componentes principales, en principio no habría razones para descartar variables del estudio, en especial, si muchos de estos índices están correlacionados entre sí.

A continuación, en la Tabla 3.1 se enumeran cada una de las variables consideradas junto con la fórmula que indica cómo se obtiene cada una de ellas. A los efectos de lograr una mejor representación de los índices seleccionados, siguiendo el criterio



adoptado por la mayoría de los autores al comentar el tema, éstos se han clasificado a priori en 6 grupos: (1) Liquidez; (2) Operativos; (3) Endeudamiento; (4) Cobertura; (5) Rentabilidad; (6) Evolución.

Índices de Liquidez

X1. Inmovilización de Activos	Activo No Corriente / Total Activo
X2. Fondo de Maniobra	Capital de Trabajo / Total Activo
X3. Liquidez Corriente	Activo Corriente / Pasivo Corriente
X4. Liquidez Seca	(Activo Corriente - Bienes de Cambio) / Pasivo Corriente
X5. Liquidez Absoluta	Disponibilidades / Pasivo Corriente

Índices Operativos

X6. Período Promedio de Cobranza	Créditos por Ventas / Ventas
X7. Período Promedio de Pago	Deudas Comerciales / Compras
X8. Período de Inventario	Bienes de Cambio Promedio / Costo de Ventas
X9. Ciclo del Efectivo	Período de Inventario + Período Prom. de Cobranza - Período Prom. de Pago
X10. Rotación de Activos	Ventas / Total Activo
X11. Rotación de Activos Corrientes	Ventas / Activo Corriente
X12. Rotación de Activos Fijos	Activo No Corriente / Ventas
X13. Rotación del Patrimonio Neto	Patrimonio Neto / Ventas
X14. Apalancamiento Operativo	(%Variación de la Utilidad Operativa) / (%Variación de las Ventas)

Índices de Endeudamiento

X15. Endeudamiento Total	Total Pasivo / Total Activo
X16. Endeudamiento Corriente	Pasivo Corriente / Total Activo
X17. Financiación Externa	Total Pasivo / Ventas
X18. Capitalización Externa	Pasivo No Corriente / (Pasivo No Corriente + Patrimonio Neto)
X19. Apalancamiento Externo	Patrimonio Neto / Total Pasivo

Índices de Cobertura

X20. Cobertura de la deuda	Utilidad Operativa / Total Pasivo
X21. Costo Financiero	Intereses / Total Pasivo

Índices de Rentabilidad

X22. Margen Bruto	Utilidad Bruta / Ventas
X23. Margen Operativo	Utilidad Operativa / Ventas
X24. Margen Neto	Utilidad Neta / Ventas
X25. Rentabilidad del Patrimonio Neto (ROE)	Utilidad Neta / Patrimonio Neto
X26. Rentabilidad Comparada	(ROE - ROE sector) / ROE sector
X27. Rentabilidad del Capital Empleado (ROCE)	Utilidad Neta / (Pasivo No Corriente + Patrimonio Neto)
X28. Rentabilidad Operativa del Capital Empleado	Utilidad Operativa / (Pasivo No Corriente + Patrimonio Neto)
X29. Rentabilidad del Activo	Utilidad Neta / Total Activo
X30. Rentabilidad Operativa del Activo	Utilidad Operativa / Total Activo

Índices de Evolución

X31. Variación de Ventas	$(Ventas_j - Ventas_{j-1}) / Ventas_{j-1}$
X32. Variación del Activo	$(Total\ Activo_j - Total\ Activo_{j-1}) / Total\ Activo_{j-1}$
X33. Variación de la Utilidad Neta	$(Utilidad\ Neta_j - Utilidad\ Neta_{j-1}) / Utilidad\ Neta_{j-1}$
X34. Variación del Pasivo	$(Total\ Pasivo_j - Total\ Pasivo_{j-1}) / Total\ Pasivo_{j-1}$
X35. Variación de la Inmovilización de Activos	$(Inmov.\ de\ Activos_j - Inmov.\ de\ Activos_{j-1}) / Inmov.\ de\ Activos_{j-1}$
X36. Variación del Fondo de Maniobra	$(Fondo\ de\ Maniobra_j - Fondo\ de\ Maniobra_{j-1}) / Fondo\ de\ Maniobra_{j-1}$
X37. Variación de la Liquidez Corriente	$(Liquidez\ Corriente_j - Liquidez\ Corriente_{j-1}) / Liquidez\ Corriente_{j-1}$
X38. Variación de la Rotación de Activos	$(Rotación\ de\ Activos_j - Rotación\ de\ Activos_{j-1}) / Rotación\ de\ Activos_{j-1}$
X39. Variación del Endeudamiento Total	$(Endeudamiento\ Total_j - Endeudamiento\ Total_{j-1}) / Endeudamiento\ Total_{j-1}$
X40. Variación de la Financiación Externa	$(Financiación\ Externa_j - Financiación\ Externa_{j-1}) / Financiación\ Externa_{j-1}$
X41. Variación de la Cobertura de la deuda	$(Cobertura\ de\ la\ deuda_j - Cobertura\ de\ la\ deuda_{j-1}) / Cobertura\ de\ la\ deuda_{j-1}$
X42. Variación del Margen Operativo	$(Margen\ Operativo_j - Margen\ Operativo_{j-1}) / Margen\ Operativo_{j-1}$
X43. Variación del Margen Neto	$(Margen\ Neto_j - Margen\ Neto_{j-1}) / Margen\ Neto_{j-1}$
X44. Variación de la Rentabilidad del Patrimonio Neto	$(ROE_j - ROE_{j-1}) / ROE_{j-1}$
X45. Variación de la Rentabilidad del Activo	$(Rent.\ del\ Activo_j - Rent.\ del\ Activo_{j-1}) / Rent.\ del\ Activo_{j-1}$

Tabla 3.1 – Variables consideradas en el estudio



4. ANÁLISIS DEL PERFIL DE LA MUESTRA

Hasta aquí se han definido las empresas que formarán parte del estudio y las variables a observar en cada una de ellas. Antes de continuar con el resto del análisis, se ha decidido mostrar en esta sección el comportamiento que tienen las variables más representativas en la totalidad de la muestra. De esta manera se pretende contar con una visión global de cómo se distribuyen las empresas según los valores de estos índices financieros y verificar si a través de ellos es posible establecer una cierta diferenciación entre las empresas que se presentaron en concurso preventivo y las que no lo han hecho.

El procedimiento utilizado para evaluar el perfil de cada variable es el siguiente:

- (1) Se divide la muestra en dos según la pertenencia de cada empresa al grupo 1 (sin problemas financieros) o al grupo 2 (con problemas financieros).
- (2) Para cada uno de estos grupos se construye el histograma respectivo y luego se representan en un mismo gráfico.
- (3) Se comparan los histogramas teniendo en cuenta la dispersión y los promedios que toma la variable en cada grupo y se verifica el grado de diferenciación entre las empresas.

Debemos considerar que para poder representar los histogramas de cada grupo de empresas en un mismo gráfico ambos deben contar con la misma cantidad de intervalos y estar definidos en el mismo rango de la variable. Por este motivo, se ha decidido que todos los histogramas tengan un total de 7 intervalos adoptando el criterio de Sturges [4] según el cual:

$$k = 1 + \frac{\ln(n)}{\ln(2)} \quad (4.1)$$

donde k es el número de intervalos y n es el número de observaciones en la muestra.

En primer lugar analizaremos los histogramas correspondientes a la variable X_3 – *Liquidez corriente*. En la Figura 4.1 se puede observar que los dos grupos muestran una distribución similar pero con un desplazamiento de los valores medios que permite cierta distinción entre estos tipos de empresas. Por un lado, se observa que la mayoría de las empresas del grupo 1 muestran un índice de liquidez corriente mayor a la unidad, mientras que la mayoría de las empresas del grupo 2 tienen un índice menor o igual que la unidad. Este comportamiento es razonable si se tiene en cuenta que las empresas con problemas financieros son más propensas a valores bajos de liquidez.

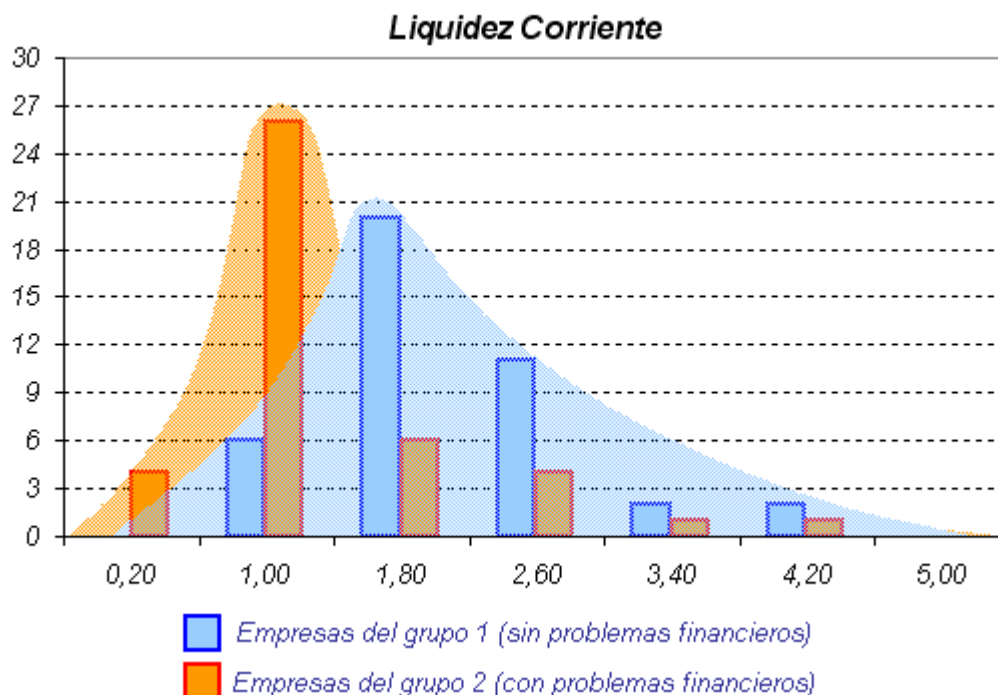


Figura 4.1 – Histogramas de la liquidez corriente

A continuación, en la Figura 4.2 se pueden observar los histogramas correspondientes a la variable X_{10} – *Rotación de activos*. Al igual que en el caso anterior, los dos grupos de empresas muestran una distribución similar pero con un cierto desplazamiento. Aquellas que tienen problemas financieros muestran una tendencia hacia valores de este índice menores que la unidad, mientras que en el resto se observa una preponderancia de valores mayores que la unidad. Esto sugiere que las empresas que no se presentaron en concurso preventivo son aquellas que en principio tienen una mayor eficiencia en la generación de ingresos, como era de esperarse.

En la Figura 4.3 se han representado los histogramas de la variable X_{14} – *Apalancamiento operativo*. En un primer análisis, algunos podrían pensar que la mayoría de las empresas con problemas financieros deberían tener un alto riesgo operativo. Sin embargo, la distribución de este índice en cada grupo es muy similar y la mayoría de las empresas en su conjunto muestran valores menores o iguales que 5. Este comportamiento sugiere que el apalancamiento operativo por sí solo no permite sacar conclusiones acerca de si una empresa pertenece a un grupo o al otro. De todos modos, en la próxima sección veremos que este índice combinado con otras variables contribuye a explicar ciertos comportamientos de las empresas.

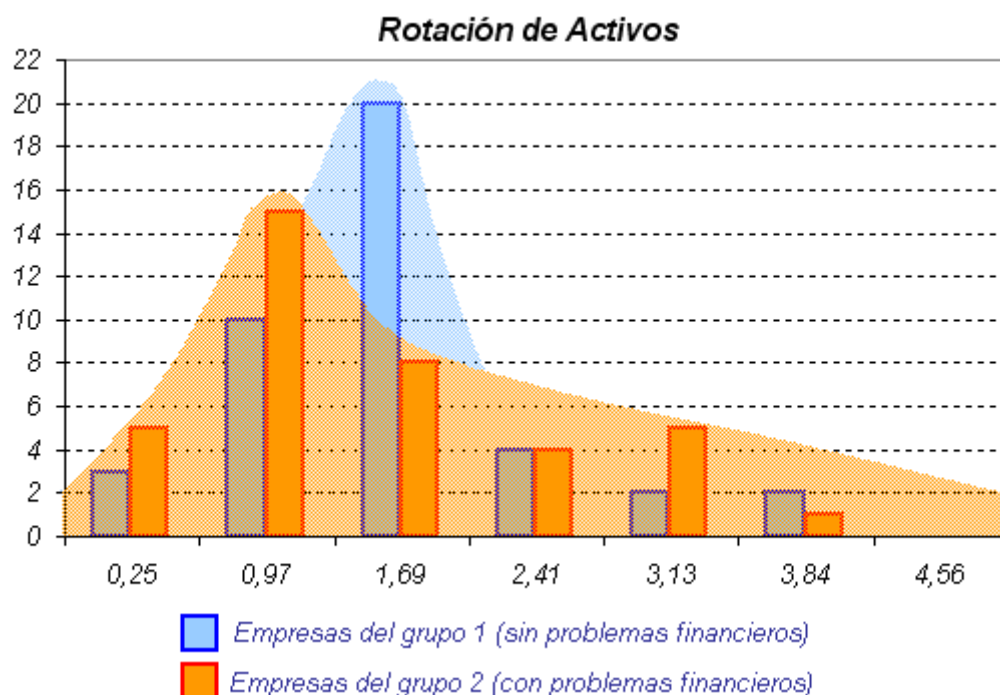


Figura 4.2 – Histogramas de la rotación de activos

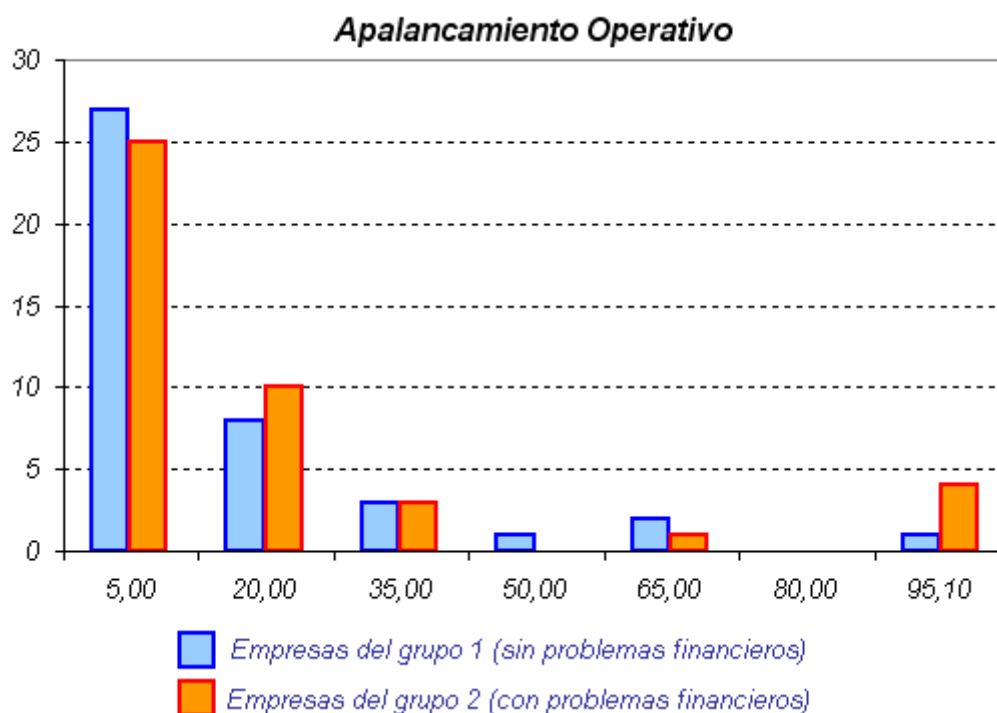


Figura 4.3 – Histogramas del apalancamiento operativo

En la Figura 4.4 se observan los histogramas de la variable X_{15} – Endeudamiento total. Como es de esperar, las empresas que se presentaron en concurso preventivo son aquellas que tienen una tendencia hacia valores altos del índice de endeudamiento. A su vez, es interesante observar que las distribuciones no son muy similares y que las



empresas del grupo 1 tienen aparentemente una distribución bastante uniforme. De todas formas, está claro que si el índice de endeudamiento es mayor a 0.7 hay aproximadamente un 74% de probabilidad de que la empresa se haya declarado en cesación de pagos.

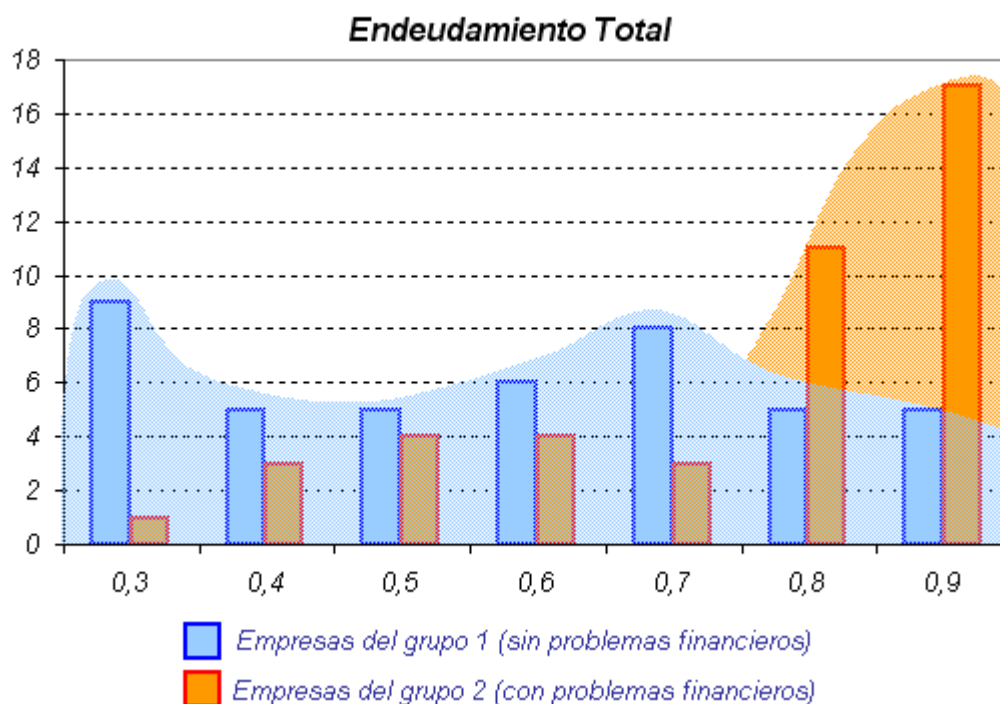


Figura 4.4 – Histogramas del endeudamiento total

En la Figura 4.5 se muestran los histogramas de la variable X_{20} – Cobertura de la deuda. En primer lugar podría decirse que cada grupo de empresas muestra una distribución similar pero con un cierto desplazamiento. Se puede observar que la mayoría de las empresas con problemas financieros tiene un índice de cobertura menor o igual que 0.2, mientras que un gran porcentaje de las empresas del grupo 1 tiene un índice de cobertura cercano a 0.3. Al mismo tiempo, hay una mayor proporción de empresas que se presentaron en concurso preventivo y que tienen una baja cobertura de la deuda acompañada de una pérdida operativa.

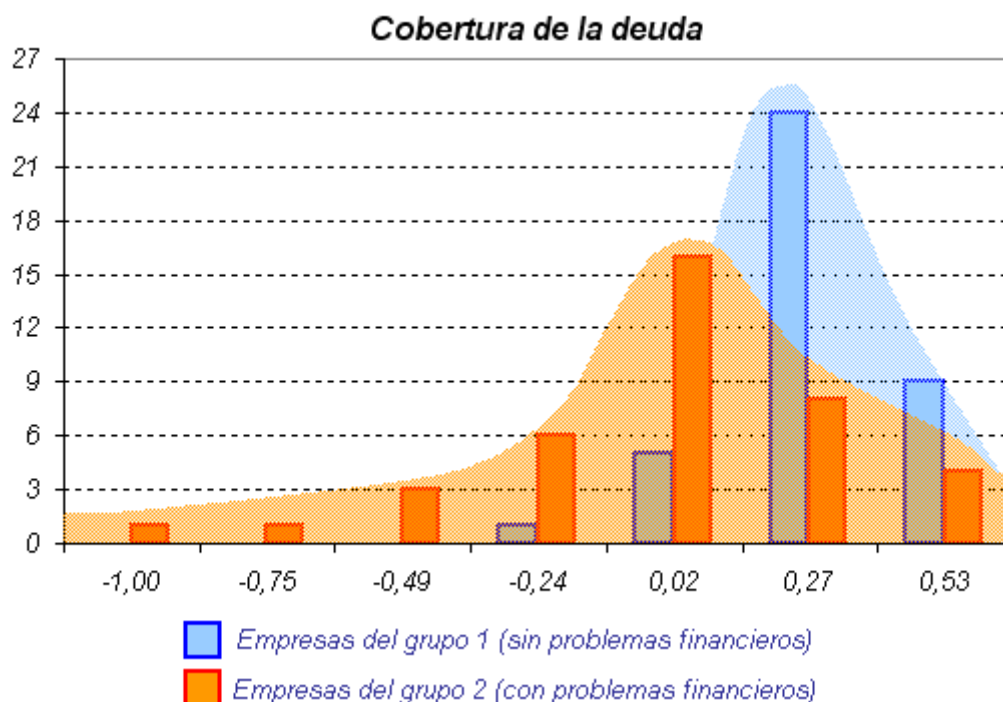


Figura 4.5 – Histogramas de la cobertura de la deuda

Por último, en la Figura 4.6 se representan los histogramas de la variable X_{29} – *Rentabilidad del activo*. A primera vista, se podría decir que cada grupo de empresas tiene una distribución distinta. Aquellas que pertenecen al grupo 1 tienen una tendencia hacia valores positivos de la rentabilidad del activo y están muy concentradas en valores cercanos a una rentabilidad del 11%, es decir, no hay una gran dispersión de los datos. En cambio, si analizamos a las empresas del grupo 2, se observa que tienen una tendencia hacia valores negativos y que la mayoría de estas tienen una rentabilidad menor o igual a un -10%, aunque en este caso debemos aclarar que la dispersión es mucho mayor.

En el gráfico es difícil notar esta dispersión debido a la cantidad de intervalos representados, que determinan que la mayoría de los valores observados se acumulen en un único intervalo negativo. En realidad, si se tomara una mayor cantidad de intervalos, podría observarse que esa concentración aparente de empresas son pequeños grupos dispersos en el sentido negativo de las abscisas. Es importante destacar que este comportamiento se repite en la mayoría de las variables relacionadas con los márgenes y la rentabilidad. De todas formas, no hay que perder de vista que la mayoría de las empresas que se presentaron en concurso preventivo tienen una rentabilidad negativa.

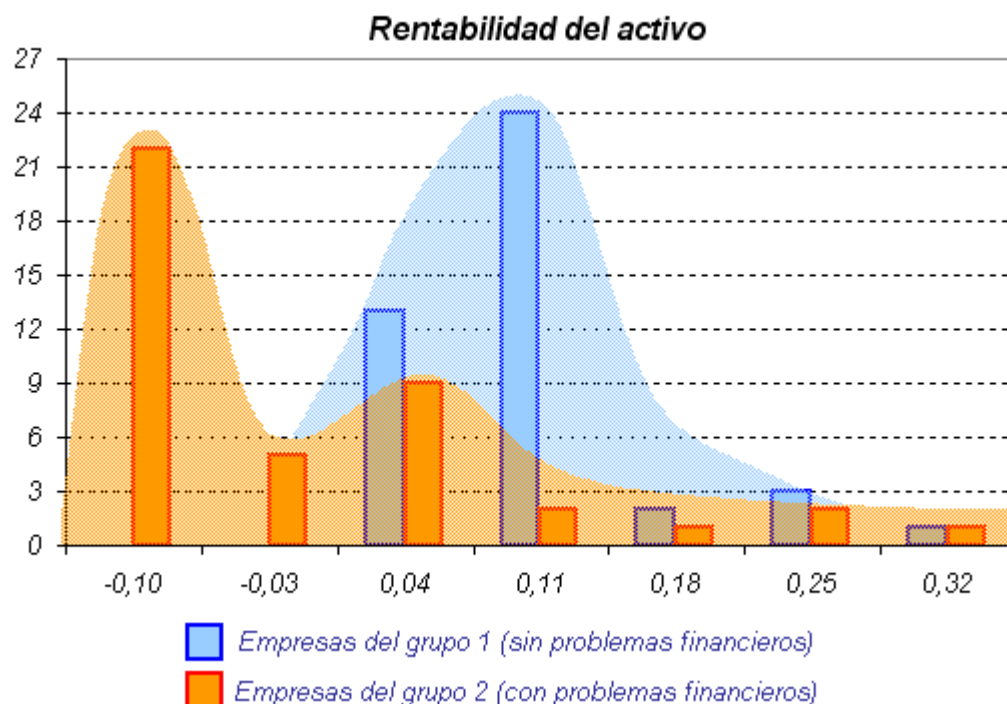


Figura 4.6 – Histogramas de la rentabilidad del activo

En resumen, hemos visto a través de los histogramas que la mayoría de los índices financieros representativos muestran un comportamiento diferente en cada uno de los grupos de empresas. En algunos casos existe un desplazamiento de los valores medios y en otros se encuentra una diferencia en la dispersión de los valores observados. Esto sugiere que existe un cierto poder de discriminación en las variables seleccionadas. En las próximas secciones veremos cómo utilizar esta información para poder identificar a las empresas con problemas financieros.



5. ANÁLISIS DE COMPONENTES PRINCIPALES

5.1. Introducción

El Análisis de Componentes Principales es un método de representación de datos que trabaja sobre un único conjunto de variables cuantitativas que describen diferentes características de un grupo de individuos. El objetivo principal que se persigue con este análisis es obtener una representación sintética de dichos individuos mediante un nuevo conjunto menor de variables que contengan prácticamente la misma información que las variables originales.

Estas nuevas variables, que justamente se denominan componentes principales, se establecen como combinaciones lineales de las variables originales y tienen la propiedad fundamental de ser incorrelacionadas entre sí y su variabilidad total es igual a la del conjunto de variables originales. En principio se definen tantas componentes principales como variables observadas de los individuos haya en la muestra. El procedimiento indica que luego deben ordenarse dichas componentes por varianzas decrecientes y se evalúa la varianza acumulada hasta que ésta alcance un porcentaje suficientemente alto de la varianza total. Este porcentaje nos indica el punto de corte a partir del cual se descartarán el resto de las componentes analizadas dado que la información que aportan, es decir su varianza respecto de la varianza total, es despreciable y por lo tanto no es posible diferenciar a los individuos a partir de ellas. De esta manera, puede obtenerse un nuevo conjunto reducido de variables capaces de representar a los individuos de la muestra.

Este método tiene sentido aplicarlo cuando las variables originales consideradas tienen cierta correlación entre sí, de modo que es posible reducir la dimensión del conjunto y obtener una mejor representación de los datos con pocas variables significativas, teniendo en cuenta por supuesto, que hay una cierta pérdida de información. De lo contrario, si las variables están poco correlacionadas, entonces la aplicación de este método no ofrecerá resultados interesantes. Por esta razón, en la selección de variables se han tenido en cuenta dentro de un mismo grupo indicadores similares, de manera de incluir en el estudio una gran cantidad de variables esperando luego poder compactar dicha información mediante este análisis.

En forma genérica, utilizaremos la variable F para referirnos a las componentes principales que pueden representarse mediante la siguiente expresión:

$$F_{ij} = V_{1j} X_{t1} + \dots + V_{ij} X_{ti} + \dots + V_{kj} X_{tk} \quad (5.1)$$

donde V_{ij} es el coeficiente de la variable i correspondiente a la componente j , X_{ti} es el valor que toma la variable X_i en el individuo t , y F_{ij} es el valor que toma la



componente F_j en el individuo t . Nótese que cada componente principal F_{tj} queda expresada como una combinación lineal de los valores que toman las variables X_i en el mismo individuo t , siendo k el número total de variables y componentes consideradas.

Las componentes principales F_{tj} cumplen con las siguientes condiciones:

- (1) Tienen media nula, es decir, $\sum_{t=1}^n F_{tj} = 0$ para todo j
- (2) Son incorrelacionadas dos a dos, es decir, $\sum_{t=1}^n F_{ti} F_{tj} = 0$ para todo $i \neq j$
- (3) Sus varianzas son decrecientes, es decir, $D^2(F_1) \geq D^2(F_2) \geq \dots \geq D^2(F_k)$

5.2. Cálculo de las componentes principales

En primer lugar expresaremos el sistema de ecuaciones que supone la expresión (5.1) en forma matricial para luego explicar cómo calcular las componentes principales para un determinado conjunto de datos.

Los valores que toman las variables observadas en los individuos pueden representarse mediante una matriz de orden $(n \times k)$, siendo n el número de individuos y k el número de variables observadas:

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \quad (5.2)$$

A su vez, los coeficientes V_{ij} de las combinaciones lineales forman una matriz cuadrada de orden $(k \times k)$:

$$\mathbf{V} = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1k} \\ V_{21} & V_{22} & \cdots & V_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ V_{k1} & V_{k2} & \cdots & V_{kk} \end{bmatrix} \quad (5.3)$$



Por último, podemos representar a las componentes principales mediante la siguiente matriz de orden $(n \times k)^3$.

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1k} \\ F_{21} & F_{22} & \cdots & F_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \cdots & F_{nk} \end{bmatrix} \quad (5.4)$$

Así la expresión (5.1) en forma matricial resulta:

$$\mathbf{F} = \mathbf{X} \cdot \mathbf{V} \quad (5.5)$$

El cálculo de las componentes principales se basa en tres enunciados que se describen a continuación y cuya demostración puede encontrarse en el libro de Daniel Peña [5] o en el de Johnson Dallas [6].

Enunciado 1 – La matriz $\mathbf{V}$ de los coeficientes de las combinaciones lineales, es la matriz cuyas columnas son los autovectores normalizados de la matriz de covarianzas $\mathbf{C}$, que a su vez, se obtiene a partir de la matriz de datos $\mathbf{X}$ centrada. La matriz de covarianzas se calcula mediante la siguiente expresión:

$$\mathbf{C} = (\mathbf{X}^T \cdot \mathbf{X}) (1/n) \quad (5.6)$$

siendo n el número total de individuos observados.

Enunciado 2 – Las varianzas de las componentes principales son iguales a los autovalores de la matriz de covarianzas $\mathbf{C}$ que es definida positiva. Es decir:

$$D^2(F_i) = \lambda_i \quad (5.7)$$

donde λ_i es el autovalor i de la matriz de covarianzas. Nótese que la cantidad de autovalores coincide con la cantidad de variables originales observadas, y por ende, con la cantidad de componentes principales analizadas.

Enunciado 3 – La suma de las varianzas de las componentes principales es igual a la suma de las varianzas de las variables originales.

³ Recordemos que en una primera instancia la cantidad de componentes principales es igual a la cantidad de variables observadas.



Es importante destacar que para obtener la matriz de covarianzas se debe partir de la matriz de datos $\mathbf{X}$ centrada, es decir, a cada valor observado se le debe restar la media de la variable en cuestión. Existe la posibilidad de trabajar con los datos estandarizados, es decir, se le resta a cada valor observado la media y además se lo divide por el desvío de la variable que corresponda. En tal caso, mediante la expresión (5.6) se obtiene la matriz de correlaciones $\mathbf{R}$ y el procedimiento planteado es igualmente válido pero los resultados que se obtienen son diferentes. Esto último implica que el análisis de componentes principales es sensible a cambios de escala, siendo éste uno de sus principales puntos débiles como método de representación de datos. Sin embargo, en general cuando los valores de las variables originales son similares entre sí, es conveniente trabajar con los datos centrados únicamente. Por este motivo, se ha elegido trabajar con la matriz de covarianzas en nuestro estudio, considerando que la mayoría de las variables son índices financieros con rangos de variabilidad similares.

Con el objetivo de interpretar el significado que pueden tener las componentes principales, es necesario calcular las correlaciones que existen entre cada variable original y dichas componentes. En forma genérica, la correlación entre la variable X_i y la componente F_j puede calcularse de la siguiente manera:

$$R_{ij} = \frac{\text{Cov}(X_i, F_j)}{D(X_i)D(F_j)} \quad (5.8)$$

y de acuerdo con los enunciados mencionados anteriormente, resulta:

$$R_{ij} = \frac{V_{ij}\sqrt{\lambda_j}}{S_i} \quad (5.9)$$

donde R_{ij} representa la correlación entre la variable i y la componente j , V_{ij} es el correspondiente elemento de la matriz de autovectores $\mathbf{V}$, λ_j es el autovalor de la componente principal y S_i es el desvío de la variable en cuestión.

A modo de síntesis, a continuación se muestran 6 pasos a seguir para poder aplicar el análisis de componentes principales.

Paso 1: Centrar la matriz de datos $\mathbf{X}$ y calcular la matriz de covarianzas $\mathbf{C}$ según la expresión (5.6).

Paso 2: Calcular los autovalores de la matriz de covarianzas $\mathbf{C}$. De acuerdo con lo expuesto en el Enunciado 2, dichos autovalores son iguales a las varianzas de las



componentes principales. Por tal motivo, se deben ordenar los autovalores en orden decreciente y luego calcular la varianza acumulada total con cada componente⁴.

Paso 3: Seleccionar como variables representativas de los datos únicamente las primeras componentes principales que acumulen aproximadamente un 90% de la varianza total y descartar el resto.

Paso 4: Hallar la matriz **V** mediante el cálculo de los autovectores de la matriz de covarianzas **C** según el Enunciado 1.

Paso 5: Calcular la matriz de componentes principales **F** mediante la expresión (5.5). De ser posible, representar en un gráfico F_i vs. F_j los valores que toman las componentes principales seleccionadas en la muestra de manera de contar con una visualización de la distribución de los individuos.

Paso 6: Calcular para todas las variables originales y las componentes principales las correlaciones según la expresión (5.9). De ser posible, adjudicar una interpretación a las componentes principales seleccionadas teniendo en cuenta las variables originales sobre las cuales cargan significativamente⁵.

5.3. Resultados obtenidos

En este apartado comentaremos los resultados que se obtuvieron al aplicar el análisis de componentes principales en la muestra de datos compuesta por 86 empresas y con un total de 45 observaciones sobre cada una de ellas según se detalló en las secciones 2 y 3.

En primer lugar, se calcularon los autovalores de la matriz de covarianzas **C** y se comprobó que los dos primeros abarcan un 93% de la varianza total (ver Anexo 10.4). Por este motivo, se decidió retener las dos primeras componentes principales F_1 y F_2 para representar los datos de la muestra. Éstos resultados son significativos teniendo en cuenta que se logró reducir el espacio de representación a dos dimensiones partiendo de un total de 45 variables.

Posteriormente, se determinaron los valores que toman las componentes principales en cada una de las empresas de la muestra (ver Anexo 10.5) mediante el cálculo previo de la matriz de autovectores **V**. De esta manera, se pudo representar la distribución de éstas mediante un gráfico que tiene como abscisas y ordenadas a los

⁴ Cabe aclarar que el software *Microsoft Excel* permite la instalación de un complemento gratuito denominado *Matrix* a través del cual es posible calcular los autovalores y autovectores de una matriz.

⁵ Se dice que una componente “carga” significativamente sobre una variable cuando tiene una alta correlación con ella.



valores de las componentes principales retenidas, tal como se observa en la Figura 5.1.

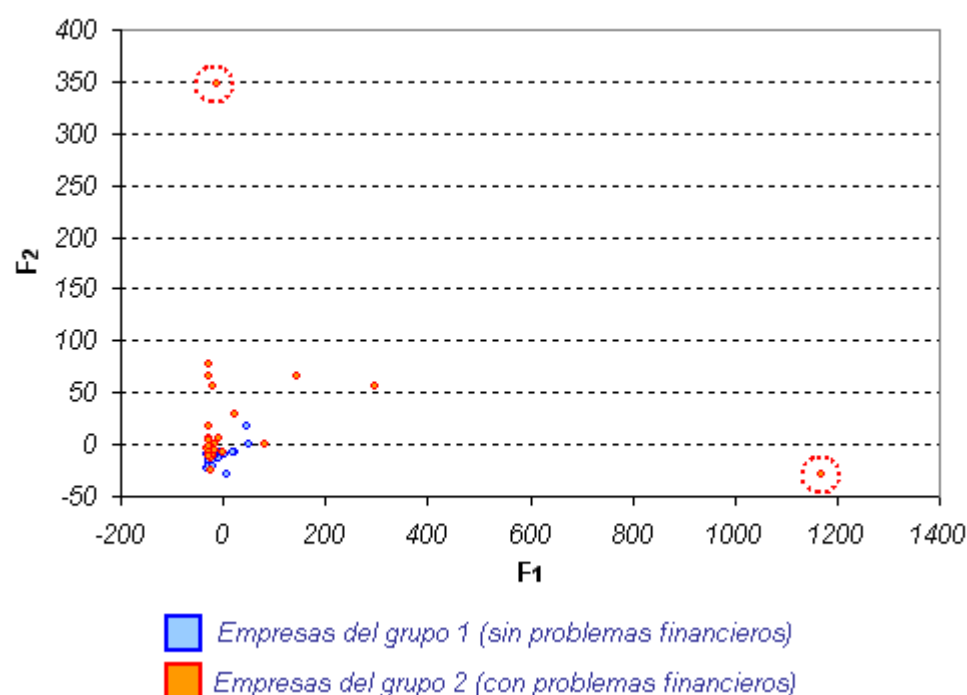


Figura 5.1 – Primera representación de las componentes principales en la muestra

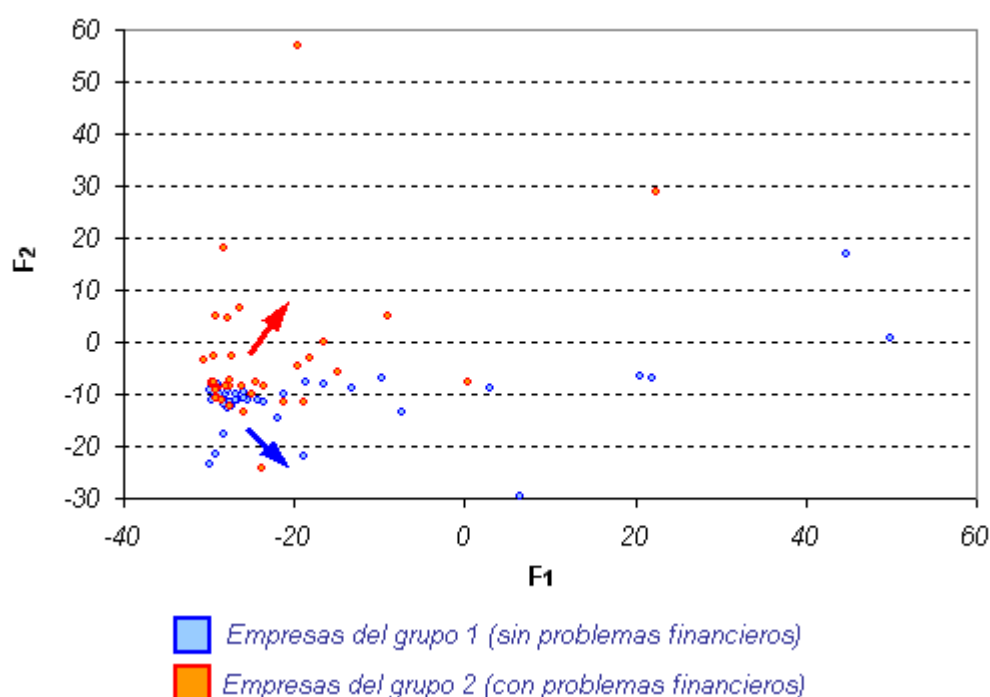


Figura 5.2 – Segunda representación de las componentes principales en la muestra



En color azul se han representado a las empresas del grupo 1 (sin problemas financieros) y en color rojo a las empresas del grupo 2 (con problemas financieros). A primera vista, en el grupo 2 se identifican mediante un círculo dos *outliers*, es decir, individuos con características particulares que se alejan demasiado del resto. Por otro lado, también es razonable esperar que al observar la distribución de dichas empresas se puedan diferenciar los dos grupos mencionados. Sin embargo, esto no es posible corroborarlo en la Figura 5.1 donde los *outliers* determinan una escala de representación que condensa la mayoría de los individuos en una pequeña área del gráfico. Por tal motivo, se decidió efectuar una nueva representación de las empresas, como se observa en la Figura 5.2, pero en este caso se excluyeron aquellas observaciones que tenían valores de las componentes principales muy distintos del resto.

A partir de este nuevo gráfico, es posible evaluar en que grado se diferencian las empresas unas de otras. Si bien no hay una clara diferenciación entre ambos grupos, aparentemente las empresas del grupo 2 muestran una tendencia hacia valores altos de F_2 mientras que en las empresas del grupo 1 ocurre lo contrario. Por otro lado, se observa una gran concentración de empresas con un valor similar y negativo de la componente F_1 , así como también individuos dispersos de ambos grupos que poseen valores más altos de dicha componente.

Por último, se calcularon las correlaciones entre las variables originales y las componentes principales según la expresión (5.9).

Los resultados demuestran que la primer componente principal F_1 tiene una alta correlación positiva con las variables X_{14} – *Apalancamiento operativo*, X_{41} – *Variación de la cobertura de la deuda* y X_{42} – *Variación del margen operativo* (ver Anexo 10.6). Esto sugiere que dicha componente refleja el riesgo operativo y la evolución en la capacidad de generar utilidades operativas por parte de las empresas. De esta manera, se podría denominar a esta nueva variable como *Variación Operativa*.

Un valor alto de esta componente puede ser causa de: (1) un alto apalancamiento operativo; (2) una mejora de la cobertura de la deuda; (3) una mejora del margen operativo; (4) una combinación de las alternativas mencionadas. De todas formas, se puede verificar a través de la matriz de autovectores que el apalancamiento operativo es la variable que posee mayor influencia sobre F_1 .

Por lo tanto, en principio se podría pensar que aquellas empresas que exhiben valores altos de esta componente presentan un elevado apalancamiento, respaldado posiblemente con una mejora de la cobertura de la deuda y del margen operativo. De hecho, si se observa en la Figura 5.2, las empresas que se alejan hacia valores altos de F_1 y mantienen un valor de la componente F_2 similar al promedio de la muestra, presentan estas características. A su vez, hay que destacar a las empresas que si bien presentan valores altos de F_1 también presentan valores altos de F_2 . Al analizar en



detalle dichas observaciones se pudo comprobar que justamente estas empresas tienen una fuerte caída, tanto de la cobertura de la deuda como del margen operativo, y el valor tan elevado de F_1 se debe exclusivamente a un valor muy alto del apalancamiento operativo.

En conclusión, esto último sugiere que las empresas con una alta variación operativa sólo muestran una mejora de la cobertura de la deuda y del margen operativo cuando mantienen un valor de F_2 similar o menor al promedio de la muestra. Por otro lado, aquellas empresas que muestran valores altos de ambas componentes son aquellas que tuvieron una alta variación de sus operaciones con una pérdida de la cobertura de la deuda y del margen operativo. Por lo tanto, es de esperar que las empresas con problemas financieros tengan estas características, aunque esto no sea una condición suficiente. Es decir, si una empresa ha sufrido una variación operativa negativa, no necesariamente debe presentar problemas financieros.

En la Figura 5.3 se puede observar una banda de color amarillo que representa a un grupo numeroso de empresas que no han sufrido una gran variación operativa y un sector de color verde donde se encuentran algunas empresas dispersas de mayor variación operativa. Teniendo en cuenta que tanto las empresas del grupo 1 como las del grupo 2 tienen valores altos y bajos de esta componente, resulta difícil diferenciarlas a partir de esta variable únicamente. Sin embargo, cuando esta información se complementa con la componente F_2 se pueden diferenciar ciertas características de las empresas. A continuación, veremos la interpretación de la componente F_2 .

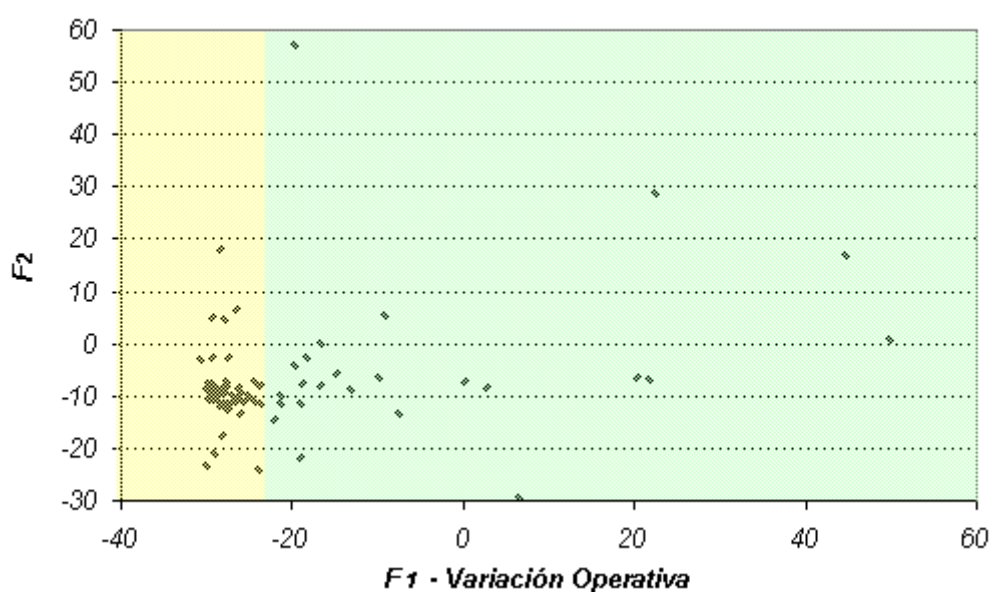


Figura 5.3 – Sectorización de la muestra según la variación operativa



Al evaluar las correlaciones entre las variables originales y la componente principal F_2 , se observa que ésta posee una alta correlación negativa con las variables X_{33} – Variación de la utilidad neta, X_{43} – Variación del margen neto y X_{45} – Variación de la rentabilidad del activo (ver Anexo 10.6). Por lo tanto, esta componente estaría reflejando en mayor medida la evolución de la capacidad y de la eficiencia de las empresas para generar ganancias. Es razonable entonces denominar a esta variable como *Variación de la Rentabilidad*.

Un valor alto de esta componente puede ser causa de: (1) una disminución de la utilidad neta; (2) una disminución del margen neto; (3) una disminución de la rentabilidad del activo; (4) una combinación de las alternativas mencionadas. Esto determina que las empresas con un alto valor de esta componente son aquellas que habrían sufrido un considerable deterioro de su rentabilidad. Tal es así que, si observamos la Figura 5.2, se puede corroborar que la mayoría de las empresas que presentan valores elevados de F_2 pertenecen al grupo 2, es decir, que se han presentado en concurso preventivo. Por otro lado, se pueden observar algunas pocas empresas que si bien pertenecen a dicho grupo muestran valores bajos de esta componente. En estos casos, se debería interpretar que dichas empresas se están recuperando de sus problemas financieros a través de una mejora de su rentabilidad.

En la Figura 5.4 se puede observar un sector en color rojo que representa a las empresas que han tenido un deterioro considerable de su rentabilidad, y en oposición, hay un sector en color verde que corresponde a aquellas que lograron una mejora de su rentabilidad. A su vez, se representa un sector en color amarillo que corresponde a aquellas empresas con una evolución de la rentabilidad similar y cercana al promedio de la muestra.

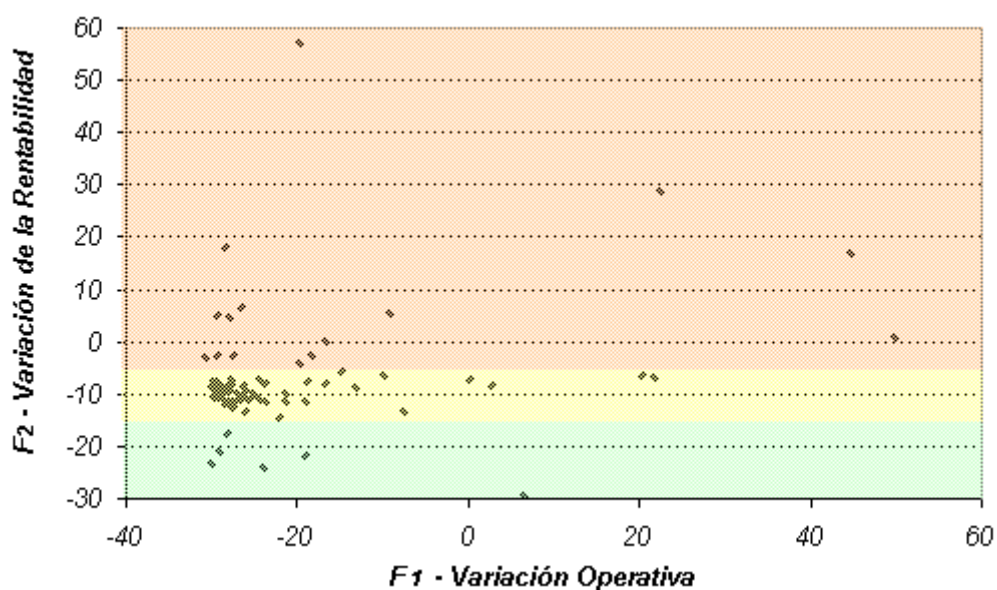


Figura 5.4 – Sectorización de la muestra según la variación de la rentabilidad



Si ahora efectuamos un análisis simultáneo de las dos componentes principales que se han retenido, es posible representar de manera aproximada y en un único gráfico, 6 sectores que definen distintas variaciones en las empresas tal como se muestra en la Figura 5.5. Al pie de dicha figura se detallan las evoluciones que han tenido las empresas que pertenecen a cada sector.

Los sectores definidos son sólo una referencia que nos permite reconocer el comportamiento que ha mostrado una empresa en los últimos ejercicios desde el punto de vista de su operación y su rentabilidad. En principio, esperaríamos que las empresas con mayor predisposición hacia la cesación de pagos se ubiquen en los sectores 1 y 2 que representan un deterioro operativo y/o de la rentabilidad. Por otro lado, las empresas que no han tenido problemas financieros generalmente deberían estar en los sectores 5 y 6 que se identifican con signos de estabilidad y/o mejoras en la operación y la rentabilidad. Respecto de los sectores 3 y 4 no es posible asociarlos a alguno de los grupos de empresas. De todas maneras, hay que dejar en claro que si bien el gráfico es útil para reconocer la evolución de una empresa, no siempre dicha evolución determina si una empresa tiene problemas financieros.

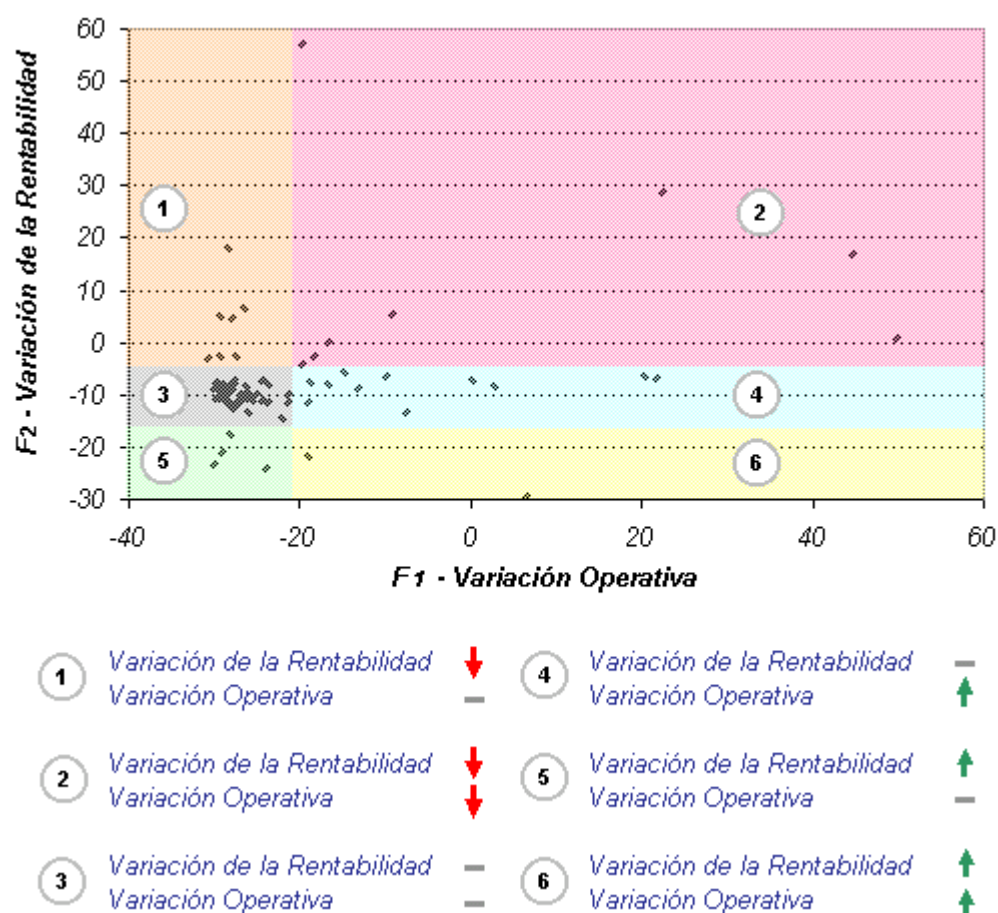


Figura 5.5 – Sectorización de la muestra según la variación operativa y de la rentabilidad



En resumen, los resultados obtenidos al aplicar el análisis de componentes principales demuestran que esta técnica ha sido de gran ayuda para lograr una mejor representación de los individuos de la muestra fundamentalmente debido al poder de síntesis que brinda.

Al mismo tiempo, hemos podido identificar distintos sectores que responden a ciertas variaciones operativas y de la rentabilidad de las empresas a través de las cuales es posible evaluar la predisposición que poseen hacia los problemas financieros. En la próxima sección, con el fin de mejorar la precisión en la clasificación de las empresas se aplicará el análisis de regresión logística para desarrollar un modelo que nos permita estimar la probabilidad de que una empresa se presente en concurso preventivo.





6. ANÁLISIS DE REGRESIÓN LOGÍSTICA

6.1. El modelo de regresión logística

Dentro de los métodos estadísticos multivariantes, el análisis de regresión lineal es uno de los más utilizados cuando se intenta describir la relación entre una variable dependiente o explicada y un conjunto de variables independientes o explicativas. El objetivo que se persigue es predecir el comportamiento de la variable explicada a partir de las variables explicativas.

Sin embargo, existen otros tipos de análisis quizá menos conocidos que también pueden utilizarse para la búsqueda de modelos estadísticos, como por ejemplo, la regresión logística que se aplica en el presente estudio. El objetivo que se busca en este caso es estimar la probabilidad de ocurrencia de un suceso en función de un conjunto de variables causales del mismo.

La principal diferencia entre estos dos análisis, es que en el caso de la regresión lineal la variable explicada debe ser continua mientras que en la regresión logística la variable explicada, en la gran mayoría de los casos, es binaria o dicotómica. Esto supone la consideración de hipótesis y modelos paramétricos distintos, pero en general el proceso de análisis sigue siendo el mismo.

Justamente el hecho de que la variable explicada sea dicotómica es la razón por la cual se ha elegido utilizar la regresión logística en este estudio. El interés radica en identificar a las empresas que tienen problemas financieros a partir de determinadas características observadas que son útiles para estimar la probabilidad de que se presenten en concurso preventivo en el corto plazo.

En cualquier modelo de regresión, una de las principales características a determinar es el valor esperado de la variable explicada dado un conjunto de variables independientes. Éste se denota mediante $E(Y / \bar{X})$ donde Y es la variable explicada y $\bar{X}$ es el vector de las variables explicativas.

Es común encontrar en la bibliografía que se utilice $\pi(\bar{X}) = E(Y / \bar{X})$ cuando se refiere al modelo de regresión logística para simplificar la notación. Dicho valor está dado por la siguiente expresión:

$$\pi(\bar{X}) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}} \quad (6.1)$$



donde $\beta_0, \beta_1, \dots, \beta_k$ representan los parámetros a estimar del modelo. Es importante destacar que si bien $\pi(\bar{X})$ representa el valor esperado de la variable dependiente Y , los valores que éste toma son también las probabilidades de ocurrencia del suceso caracterizado por dicha variable.

Dado que la variable de salida del modelo es dicotómica, debemos exigir que el dominio de $E(Y / \bar{X})$ esté comprendido entre 0 y 1 para cualquier rango de valores de las variables explicativas. Esta exigencia determina una diferencia respecto del modelo de regresión lineal donde la cantidad $E(Y / \bar{X})$ puede tomar cualquier valor entre $-\infty$ y $+\infty$ dependiendo del rango de valores de las variables explicativas.

Por esta razón, en muchas ocasiones se utiliza una transformación de $\pi(\bar{X})$ denominada función *logit* que se define como:

$$g(\bar{X}) = \text{Ln} \left[\frac{\pi(\bar{X})}{1 - \pi(\bar{X})} \right] \quad (6.2)$$

La importancia de esta transformación consiste en que tiene muchas de las propiedades deseables de un modelo de regresión lineal que son: (1) la linealidad en los parámetros; (2) la continuidad de la función; (3) el rango de variación de $-\infty$ hasta $+\infty$ dependiendo del rango de valores de las variables explicativas. Esto se puede ver mejor si expresamos $g(\bar{X})$ como:

$$g(\bar{X}) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (6.3)$$

En la expresión (6.2), $\pi(\bar{X})/[1 - \pi(\bar{X})]$ indica la relación entre la probabilidad de que se produzca el suceso y la de que no se produzca. Este cociente se denomina comúnmente "*chance*". Por ejemplo, si un equipo de fútbol tiene una probabilidad de 2/3 de ganar un partido, se dice que su *chance* es de 2 a 1.

En general, el valor observado de la variable explicada puede expresarse como:

$$Y = E(Y / \bar{X}) + \varepsilon \quad (6.4)$$

donde ε es la perturbación. Esta variable introduce el componente aleatorio en el modelo. Cabe aclarar que en el análisis de regresión, si bien las variables explicativas pueden ser aleatorias, éstas no se consideran como tales dado que el valor de la variable explicada se estima para un conjunto de valores *dados* de las variables explicativas.



En el modelo de regresión lineal se asume que la perturbación posee una distribución normal con media nula y desvío constante. Éste no es el caso en la regresión logística donde el valor de la variable dicotómica que se desea explicar resulta:

$$Y = \pi(\bar{X}) + \varepsilon \quad (6.5)$$

En estas condiciones, la perturbación ε puede tomar dos valores posibles. Cuando $Y = 1$, entonces $\varepsilon = 1 - \pi(\bar{X})$ con una probabilidad $\pi(\bar{X})$, o bien, cuando $Y = 0$, entonces $\varepsilon = -\pi(\bar{X})$ con probabilidad $1 - \pi(\bar{X})$. Por lo tanto, se puede decir que la perturbación ε posee una distribución con media nula y una varianza de $\pi(\bar{X})[1 - \pi(\bar{X})]$. Esto sugiere que la distribución de la perturbación y consecuentemente de la variable explicada es binomial, a diferencia de la regresión lineal donde el análisis se sustenta en una distribución normal.

6.2. Estimación de los parámetros

Existen tres métodos utilizados para la estimación de los parámetros en el modelo de regresión logística, a saber: (1) Estimación por mínimos cuadrados; (2) Estimación por mínimos cuadrados ponderados; (3) Estimación por máxima verosimilitud.

Dado que en cualquiera de estos métodos nos encontramos con un modelo no lineal, será necesario aplicar un procedimiento iterativo para optimizar una determinada función objetivo que nos permita encontrar una estimación de los parámetros. En este estudio se ha utilizado el software *Microsoft Excel* como herramienta de cálculo iterativo.

6.2.1. Estimación por mínimos cuadrados

Mediante este procedimiento se pretende encontrar los parámetros del modelo que minimicen la siguiente expresión:

$$Q = \sum_{t=1}^n [Y_t - \pi(\bar{X}_t)]^2 \quad (6.6)$$

donde Y_t son los valores observados de la variable explicada para cada individuo t y $\pi(\bar{X}_t)$ son sus valores esperados. Es decir, el objetivo es encontrar el conjunto de parámetros para los cuales la suma de los cuadrados de las diferencias entre los valores observados y los valores esperados de la variable independiente es mínima.



Si bien éste método es muy utilizado para la estimación de parámetros en la regresión lineal, en la regresión logística pueden presentarse algunos inconvenientes. En primer lugar, al buscar minimizar Q se pueden obtener valores muy elevados o demasiado pequeños de dicha sumatoria dado que es muy sensible a los cambios en los parámetros. Para evitar esto se puede minimizar la suma de los cuadrados de las diferencias ponderadas, como veremos a continuación en la estimación por mínimos cuadrados ponderados. Otra limitación radica en que no se pueden realizar los ensayos de hipótesis de nulidad de los parámetros, herramienta muy utilizada en los análisis de regresión para validar los modelos propuestos y corroborar su ajuste a la muestra de datos.

6.2.2. Estimación por mínimos cuadrados ponderados

En este caso, se buscan los parámetros del modelo que minimicen la siguiente expresión:

$$\chi^2 = \sum_{t=1}^n \frac{[Y_t - \pi(\bar{X}_t)]^2}{\pi(\bar{X}_t)[1 - \pi(\bar{X}_t)]} \quad (6.7)$$

donde χ^2 es un estadístico que posee una distribución Chi-Cuadrado con $(n-p)$ grados de libertad, siendo p el número de parámetros a estimar, bajo el supuesto de que el modelo se ajusta correctamente a los datos independientemente del método utilizado para estimar dichos parámetros.

Si el resultado de la sumatoria supera el valor crítico $\chi^2_{n-p;1-\alpha}$ correspondiente a la distribución Chi-Cuadrado, el modelo debe rechazarse. De lo contrario, el modelo puede aceptarse y entonces cuanto menor sea el resultado de la sumatoria, mejor será el ajuste del modelo. Por este motivo, el método se sustenta en hallar los parámetros que minimizan la expresión (6.7).

6.2.3. Estimación por máxima verosimilitud

Con éste método se buscan los parámetros del modelo que maximicen la probabilidad de obtener los valores observados para la muestra de datos disponible. Es decir, el método pretende encontrar los parámetros que maximizan la probabilidad de que efectivamente ocurran los sucesos observados para el conjunto de individuos de la muestra. Para ello es necesario contar con una función objetivo que relacione los parámetros del modelo con los valores observados de la muestra. Dicha función se denomina *función de verosimilitud* y en el caso particular de la regresión logística adopta la siguiente forma:



$$\zeta = \prod_{t=1}^n p_t \quad (6.8)$$

donde p_t representa la probabilidad de obtener el valor observado para cada individuo de la muestra. Dicha probabilidad según el modelo de regresión logística puede expresarse como:

$$p_t = \pi(\bar{X}_t)^{Y_t} [1 - \pi(\bar{X}_t)]^{1-Y_t} \quad (6.9)$$

Teniendo en cuenta que $p_t = \pi(\bar{X}_t)$ si $Y_t = 1$ y que $p_t = 1 - \pi(\bar{X}_t)$ si $Y_t = 0$, entonces:

$$p_t = |1 - Y_t - \pi(\bar{X}_t)| \quad (6.10)$$

Por lo tanto, la función de verosimilitud resulta:

$$\zeta = \prod_{t=1}^n |1 - Y_t - \pi(\bar{X}_t)| \quad (6.11)$$

Con el objetivo de facilitar los cálculos algebraicos, generalmente se busca maximizar el logaritmo natural de esta función. Entonces la función objetivo queda definida de la siguiente manera :

$$Ln\zeta = \sum_{t=1}^n Ln|1 - Y_t - \pi(\bar{X}_t)| \quad (6.12)$$

En resumen, los parámetros que maximizan la expresión (6.12), denominados *estimadores máximo-verosímiles*, son aquellos para los cuales la probabilidad de obtener los sucesos observados es máxima. Éste método tiene la ventaja de que permite ensayar la significancia de los parámetros, es decir, evaluar si hay suficientes pruebas estadísticas de que dichos parámetros estimados no son nulos.

6.3. Validación del modelo de regresión logística

El objetivo que se busca en este apartado es explicar en forma concisa las diferentes pruebas que pueden llevarse a cabo para validar un modelo basado en la regresión logística. Dicha validación es muy importante dado que nos permite contar con una medida cuantitativa sobre cómo el modelo planteado se ajusta a los datos de la muestra.



Siguiendo la estructura que propone Roberto M. García [7] en su libro sobre inferencia estadística, clasificaremos las pruebas de validación en tres grandes grupos, a saber: (1) Ensayos de hipótesis; (2) Medidas de bondad de ajuste; (3) Pruebas de significancia de los coeficientes de regresión.

6.3.1. Ensayos de hipótesis

A continuación, veremos dos estadísticos que permiten contrastar la siguiente hipótesis nula:

H_0) El modelo se ajusta correctamente a los datos

De acuerdo con lo que hemos visto al comentar sobre la estimación de los parámetros, suponiendo que el modelo se ajusta correctamente a los datos, el siguiente estadístico:

$$\chi^2 = \sum_{t=1}^n \frac{[Y_t - \pi(\bar{X}_t)]^2}{\pi(\bar{X}_t)[1 - \pi(\bar{X}_t)]} \quad (6.13)$$

responde a una distribución Chi-Cuadrado con $(n-p)$ grados de libertad, siendo p el número de parámetros a estimar independientemente del método utilizado para dicha estimación.

La condición de rechazo para este estadístico resulta:

$$\chi^2 \geq \chi_{n-p;1-\alpha}^2 \quad (6.14)$$

Por lo tanto, si el valor calculado del estadístico mediante la expresión (6.13) es superior al valor crítico $\chi_{n-p;1-\alpha}^2$ correspondiente a la distribución Chi-Cuadrado, la hipótesis nula debe rechazarse y esto implica que el modelo no se ajusta correctamente a los datos.

También puede utilizarse el siguiente estadístico para contrastar la misma hipótesis nula:

$$D = -2Ln\zeta \quad (6.15)$$

donde ζ es el valor máximo de la función de verosimilitud definida mediante la expresión (6.11) y D es el estadístico que se denomina *desviación*. Nuevamente, si se supone que la hipótesis nula es cierta, el estadístico tiene una distribución Chi-



Cuadrado con $(n-p)$ grados de libertad, siendo p el número de parámetros a estimar independientemente del método utilizado para dicha estimación.

En este caso, la condición de rechazo resulta:

$$D \geq \chi_{n-p;1-\alpha}^2 \quad (6.16)$$

Entonces, si la desviación es superior al valor crítico $\chi_{n-p;1-\alpha}^2$ correspondiente a la distribución Chi-Cuadrado, la hipótesis nula debe rechazarse y esto implica que el modelo no se ajusta correctamente a los datos.

Por último, veremos un estadístico que permite contrastar la siguiente hipótesis nula:

$$H_0) \beta_1 = \beta_2 = \dots = \beta_k = 0$$

Bajo el supuesto de que esta hipótesis es cierta, el siguiente estadístico:

$$G = 2(\text{Ln}\zeta - \text{Ln}\zeta_0) \quad (6.17)$$

responde a una distribución Chi-Cuadrado con k grados de libertad, siendo k el número de variables explicativas del modelo. El valor de ζ_0 es aquel que se obtiene a partir de la maximización de la función de verosimilitud anulando todos los coeficientes de las variables explicativas y estimando únicamente el parámetro β_0 , es decir, el término independiente. Se puede demostrar que dicho valor se puede obtener fácilmente mediante la siguiente expresión:

$$\zeta_0 = \left(\frac{r}{n}\right)^r \left(1 - \frac{r}{n}\right)^{n-r} \quad (6.18)$$

donde n es el número total de individuos en la muestra y r es el número de individuos para los cuales $Y_t = 1$.

Es importante destacar que a diferencia de los dos estadísticos mencionados anteriormente, en este caso se requiere que los parámetros hayan sido estimados por el método de máxima verosimilitud.

La condición de rechazo resulta:

$$G \geq \chi_{k;1-\alpha}^2 \quad (6.19)$$



Por lo tanto, si el valor calculado del estadístico mediante la expresión (6.17) es superior al valor crítico $\chi^2_{k;1-\alpha}$ correspondiente a la distribución Chi-Cuadrado, la hipótesis nula debe rechazarse y esto implica que hay pruebas suficientes como para pensar que alguno de los parámetros del modelo no es nulo. Nótese que esta hipótesis es justamente contraria a la anterior.

6.3.2. Medidas de bondad de ajuste

A continuación, se muestran algunas expresiones que permiten calcular diferentes coeficientes de determinación (R^2) con el objetivo de medir el porcentaje de la variación total de las observaciones que es explicada por el modelo.

En primer lugar, veremos como se puede obtener el R^2 “clásico” que comúnmente se utiliza en el análisis de regresión lineal. En este caso, al tratarse de una regresión logística, se cumplen las siguientes igualdades:

$$\sum_{t=1}^n Y_t = r \quad (6.20)$$

$$\sum_{t=1}^n Y_t^2 = r \quad (6.20')$$

donde r es el número de individuos para los cuales $Y_t = 1$. De esta manera, la expresión que permite calcular dicho coeficiente resulta:

$$R^2_{clásico} = 1 - \frac{Q}{r \left(1 - \frac{r}{n} \right)} \quad (6.21)$$

donde Q representa la suma de los cuadrados de las diferencias entre los valores observados y los valores esperados de la variable independiente, según se muestra en la expresión (6.6).

En segundo lugar, se puede calcular el coeficiente de determinación de McFadden mediante la siguiente expresión:

$$R^2_{McFadden} = 1 - \frac{\text{Ln} \zeta}{\text{Ln} \zeta_0} \quad (6.22)$$



donde ζ es el valor máximo que puede tomar la función de verosimilitud y ζ_0 es el valor máximo de la función de verosimilitud que se obtiene a partir de la expresión (6.18).

En tercer lugar, tenemos el coeficiente de determinación de Aldrich-Nelson, que se calcula mediante la siguiente expresión:

$$R^2_{\text{Aldrich-Nelson}} = \frac{G}{G + n} \quad (6.23)$$

donde G representa el estadístico que se obtiene a partir de la expresión (6.17).

En cuarto lugar, mediante la siguiente expresión se puede hallar el coeficiente de determinación de Cox-Snell:

$$R^2_{\text{Cox-Snell}} = 1 - e^{-\frac{G}{n}} \quad (6.24)$$

donde nuevamente G representa el estadístico que se obtiene a partir de la expresión (6.17).

Por último, podemos calcular el coeficiente de determinación de Nagelkerke:

$$R^2_{\text{Nagelkerke}} = \frac{1 - e^{-\frac{G}{n}}}{1 - e^{-\frac{2}{Ln\zeta_0}}} \quad (6.25)$$

siendo n el número total de individuos en la muestra, G el estadístico que se obtiene a partir de la expresión (6.17) y ζ_0 es el valor máximo de la función de verosimilitud que se obtiene mediante la expresión (6.18).

Es importante tener en cuenta que si bien todos los coeficientes de determinación mencionados anteriormente son alternativas propuestas por diferentes investigadores, todos ellos se calculan con el mismo objetivo de cuantificar qué tanto se ajusta el modelo a los datos de la muestra. En todos los casos, los coeficientes arrojan resultados comprendidos entre 0 y 1, indicando un ajuste más pobre cuanto menor sea dicho valor.



6.3.3. Pruebas de significancia de los coeficientes de regresión

En este apartado veremos como pueden realizarse las pruebas de significancia de los coeficientes de regresión mediante dos procedimientos denominados *Contraste de Wald* y *Contraste de Wilks*. Estas pruebas de significancia consisten en ensayar para cada parámetro del modelo la hipótesis nula de que éstos son nulos. En la medida en que dichas hipótesis sean rechazadas, entonces puede decirse que existen suficientes pruebas estadísticas para afirmar que el coeficiente no es nulo, y por lo tanto, resulta significativo. Esto último puede interpretarse como un indicador de la incidencia que posee cada una de las variables incluidas en el modelo sobre la variable dependiente.

6.3.3.1. Contraste de Wald

La hipótesis nula que se plantea es la siguiente:

$$H_0) \beta_i = 0$$

donde β_i puede ser cualquier parámetro del modelo que se está validando con $1 \leq i \leq k$.

En este contraste se utiliza un estadístico que responde aproximadamente a una distribución t de Student con $(n-k-1)$ grados de libertad, siendo k el número de variables explicativas del modelo. Su valor se obtiene a partir de la siguiente expresión:

$$t = \frac{b_i}{D(b_i)} \quad (6.26)$$

donde b_i representa la estimación del parámetro en cuestión y $D(b_i)$ su desvío.

Los desvíos $D(b_i)$ de los parámetros pueden obtenerse a partir de la matriz de covarianzas-varianzas $\mathbf{C}_v$. De hecho, los elementos de la diagonal de esta matriz son iguales a la varianza de los respectivos coeficientes de regresión. Por lo tanto, una vez calculada la matriz, automáticamente se obtienen los desvíos de los parámetros estimados.

La matriz de covarianzas-varianzas $\mathbf{C}_v$ es igual a la inversa de la matriz de información $\mathbf{W}$. Es decir,

$$\mathbf{C}_v = \mathbf{W}^{-1} \quad (6.27)$$



A su vez, la matriz de información $\mathbf{W}$ se puede calcular mediante la siguiente expresión:

$$\mathbf{W} = \mathbf{X}_a^T \cdot \mathbf{P} \cdot \mathbf{X}_a \quad (6.28)$$

donde $\mathbf{X}_a$ es la matriz de datos ampliada, $\mathbf{X}_a^T$ es su traspuesta y $\mathbf{P}$ es una matriz diagonal de orden $(n \times n)$ cuyos elementos son el producto entre la probabilidad de éxito y la probabilidad de fracaso para cada individuo de la muestra.

La matriz de datos ampliada $\mathbf{X}_a$ se define de la siguiente manera:

$$\mathbf{X}_a = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ 1 & X_{31} & X_{32} & \cdots & X_{3k} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} \quad (6.29)$$

Nótese que esta matriz es muy similar a la matriz de datos $\mathbf{X}$, pero se tiene una columna adicional con valores unitarios de manera de contemplar la varianza del término independiente de la regresión.

La matriz $\mathbf{P}$ se define de la siguiente manera:

$$\mathbf{P} = \begin{bmatrix} \pi_1(\bar{X})[1 - \pi_1(\bar{X})] & 0 & \cdots & 0 \\ 0 & \pi_i(\bar{X})[1 - \pi_i(\bar{X})] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pi_n(\bar{X})[1 - \pi_n(\bar{X})] \end{bmatrix} \quad (6.30)$$

donde $\pi_i(\bar{X})$ es la probabilidad de ocurrencia del suceso en el individuo i de la muestra.

En resumen, una vez que se obtienen los desvíos de los parámetros estimados y se calculan los estadísticos correspondientes, se pueden ensayar las hipótesis planteadas. Dichos ensayos pueden ser unilaterales o bilaterales dependiendo del sentido esperado de la variación de la variable explicada frente a una determinada variación de la variable explicativa. Por ejemplo, cuando se espera que un aumento de una determinada variable explicativa produzca un aumento o una disminución de la variable dependiente, entonces se debería efectuar un ensayo unilateral. Por el contrario, cuando existen dudas sobre los efectos que puede tener una variación de la variable explicativa sobre la variable dependiente, entonces es conveniente realizar un ensayo bilateral.



De esta manera, las condiciones de rechazo para la hipótesis nula resultan:

$$t \geq t_{n-k-1;1-\alpha} \text{ si se espera que } \beta_i > 0 \quad (6.31)$$

$$t \leq -t_{n-k-1;1-\alpha} \text{ si se espera que } \beta_i < 0 \quad (6.32)$$

$$|t| \geq t_{n-k-1;1-\alpha/2} \text{ si se espera que } \beta_i \neq 0 \quad (6.33)$$

siendo $t_{n-k-1;1-\alpha}$ y $t_{n-k-1;1-\alpha/2}$ los fractiles correspondientes de la distribución t de Student.

Por último, a través de los desvíos de los parámetros estimados pueden determinarse los correspondientes intervalos de confianza de la siguiente manera:

$$b_i \pm t_{n-k-1;1-\alpha/2} D(b_i) \quad (6.34)$$

donde nuevamente $t_{n-k-1;1-\alpha/2}$ es el fractil de la distribución t de Student.

6.3.3.2. Contraste de Wilks

Al igual que en el contraste de Wald, la hipótesis nula que se plantea es la siguiente:

$$H_0) \beta_i = 0$$

donde β_i representa cualquier parámetro del modelo que se está validando con $1 \leq i \leq k$.

La diferencia radica en el estadístico utilizado para contrastar dicha hipótesis. En este caso, dicho estadístico bajo el supuesto de que la hipótesis nula es cierta, tiene una distribución Chi-Cuadrado con 1 grado de libertad. Su valor se obtiene a partir de la siguiente expresión:

$$\chi^2 = 2(\text{Ln}\zeta - \text{Ln}\zeta_i) \quad (6.35)$$

donde ζ es el valor máximo de la función de verosimilitud definida mediante la expresión (6.11) y ζ_i es el valor máximo de la función de verosimilitud que se obtiene al eliminar el parámetro estimado i .



Teniendo en cuenta que aquí también pueden realizarse contrastes unilaterales o bilaterales, las condiciones de rechazo de la hipótesis nula resultan:

$$b_i > 0 \cap \chi^2 \geq \chi_{1;1-2\alpha}^2 \text{ si se espera que } \beta_i > 0 \quad (6.36)$$

$$b_i < 0 \cap \chi^2 \geq \chi_{1;1-2\alpha}^2 \text{ si se espera que } \beta_i < 0 \quad (6.37)$$

$$\chi^2 \geq \chi_{1;1-\alpha}^2 \text{ si se espera que } \beta_i \neq 0 \quad (6.38)$$

donde $\chi_{1;1-2\alpha}^2$ y $\chi_{1;1-\alpha}^2$ son los fractiles correspondientes de la distribución Chi-Cuadrado.

6.4. Resultados Obtenidos

6.4.1. Primer modelo a partir de las componentes principales

Teniendo en cuenta que las componentes principales F_1 – *Variación Operativa* y F_2 – *Variación de la Rentabilidad* han sido de utilidad para representar a las empresas de la muestra y que contienen un 93% de la varianza total de las 45 variables originalmente consideradas, se ha decidido utilizar estas componentes para construir un modelo de regresión logística que permita estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo.

De esta manera, se busca complementar el análisis de componentes principales con una herramienta estadística que nos permita contar con una medida más precisa a la hora de discriminar a las empresas con problemas financieros.

Cabe destacar que para desarrollar este modelo se han descartado los *outliers* detectados en el análisis de componentes principales debido a que contienen una varianza atípica que pueden distorsionar los valores observados. En la Tabla 6.1 se muestran las variables consideradas en el modelo.

<i>Símbolo</i>	<i>Nombre</i>	<i>Tipo de Variable</i>
F_1	<i>Variación Operativa</i>	Variable explicativa - Continua
F_2	<i>Variación de la Rentabilidad</i>	Variable explicativa - Continua
Y	<i>Concurso Preventivo</i>	Variable explicada - Dicotómica

Tabla 6.1 – Variables consideradas en el modelo de regresión



<i>Método de Estimación</i>	<i>Parámetro Estimado</i>		
	<i>b₀</i>	<i>b₁ (F₁)</i>	<i>b₂ (F₂)</i>
<i>Mínimos Cuadrados</i>	-0.5010	-0.1331	0.3459
<i>Mínimos Cuadrados Ponderados</i>	0.3947	-0.0082	0.0757
<i>Máxima Verosimilitud</i>	0.7616	-0.0265	0.1744

Tabla 6.2 – Resultados de la estimación de los parámetros

En primer lugar, se llevó a cabo la estimación de los parámetros mediante los tres métodos vistos anteriormente cuyos resultados se muestran en la Tabla 6.2. Debido a que algunas pruebas de validación del modelo requieren que los parámetros se hayan estimado mediante el método de máxima verosimilitud, en principio, utilizaremos esta solución como representativa del modelo. Éste último se define entonces mediante la siguiente expresión:

$$\hat{Y} = \frac{1}{1 + e^{-(0.7616 - 0.0265F_1 + 0.1744F_2)}} \quad (6.39)$$

donde $\hat{Y}$ representa la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo. Según este modelo cuanto mayor sea el valor de la componente F_1 – *Variación Operativa*, menor es la probabilidad de que una empresa se presente en concurso preventivo. Por otro lado, un valor alto de la componente F_2 – *Variación de la Rentabilidad*, determina una mayor probabilidad de que la empresa tenga problemas financieros. Recordemos que un valor alto de F_1 está dado por un alto apalancamiento operativo y un valor alto de F_2 es sinónimo de un deterioro en la rentabilidad. Por lo tanto, en principio, los signos de los parámetros estimados guardan una relación acorde a la naturaleza de las variables explicativas.

Posteriormente, se llevó a cabo la validación del modelo comenzando con los ensayos de hipótesis. Cabe destacar que en todas las pruebas de validación realizadas se ha considerado un nivel de significación del 5%⁶.

El primer ensayo realizado corresponde a la hipótesis nula: H_0) *El modelo se ajusta correctamente a los datos*. Para ello se calcularon los estadísticos correspondientes mediante las expresiones (6.13) y (6.15), siendo los resultados del ensayo los que se muestran en la Tabla 6.3.

⁶ Se llama nivel de significación en un ensayo de hipótesis a la máxima probabilidad de cometer un *Error Tipo I*, es decir, rechazar una hipótesis nula que en realidad es verdadera.



Ensayo de hipótesis	H_0) El modelo se ajusta correctamente a los datos	
Estadístico calculado	$\chi^2 = 71.4825$	$D = 93.4616$
Valor crítico	$\chi^2_{81,0.95} = 103.0095$	$\chi^2_{81,0.95} = 103.0095$
Condición de Rechazo	$\chi^2 \geq \chi^2_{81,0.95}$	$D \geq \chi^2_{81,0.95}$
Resultado	No Rechazar	No Rechazar

Tabla 6.3 – Resultados del primer ensayo global

En ninguno de los dos casos se rechaza la hipótesis nula, y en principio, podría decirse que no hay suficientes pruebas estadísticas que permitan afirmar que el modelo no se ajusta correctamente a los datos.

El próximo ensayo realizado corresponde a la hipótesis nula: H_0) $\beta_1 = \beta_2 = \dots = \beta_k = 0$. En este caso, el estadístico se calculó mediante la expresión (6.17), siendo los resultados del ensayo los que se muestran en la Tabla 6.4.

Ensayo de hipótesis	H_0) $\beta_1 = \beta_2 = \dots = \beta_k = 0$
Estadístico calculado	$G = 22.9395$
Valor crítico	$\chi^2_{20,95} = 5.9915$
Condición de Rechazo	$G \geq \chi^2_{20,95}$
Resultado	Rechazar

Tabla 6.4 – Resultados del segundo ensayo global

Dado que la hipótesis nula se rechaza, en consecuencia, alguno de los parámetros estimados no es nulo.

A continuación, se calcularon las medidas de bondad de ajuste que consisten en los diferentes coeficientes de determinación. En la Tabla 6.5 se muestra un resumen de los valores obtenidos.

En principio, los valores de los R^2 sugieren que el modelo explica un porcentaje reducido de la varianza total de la variable dependiente, siendo el valor máximo de todos los coeficientes calculados $R^2_{Nagelkerke} = 0.3187$. De todas formas, no hay que perder de vista que estas medidas son referenciales y si bien permiten comparar dos modelos entre sí, en mi opinión no deben considerarse como una condición suficiente para rechazar un modelo. De hecho, a continuación evaluaremos la capacidad de discriminación del modelo al aplicarlo sobre la muestra de datos, para contar con una



medida indicativa de la eficacia que posee al clasificar a una empresa en un grupo u otro.

<i>Medidas de bondad de ajuste</i>	<i>Valor</i>
$R^2_{Clásico}$	0.2514
$R^2_{McFadden}$	0.1971
$R^2_{Aldrich-Nelson}$	0.2145
$R^2_{Cox-Stell}$	0.2390
$R^2_{Nagelkerke}$	0.3187

Tabla 6.5 – Valores obtenidos de los coeficientes de determinación

En la Tabla 6.6 se resume el porcentaje de casos en los que el modelo ha clasificado correctamente a una empresa dentro de alguno de los dos grupos, así como aquellos casos en los cuales la clasificación fue incorrecta. Para ello se utilizó el criterio de que las empresas cuya probabilidad de presentarse en concurso preventivo sea menor o igual a 0.5 se consideran dentro del grupo 1, y en caso contrario, se consideran dentro del grupo 2.

	<i>Clasificaciones Correctas</i>	<i>Clasificaciones Incorrectas</i>	<i>Total</i>
Grupo 1	95.35%	4.65%	100%
Grupo 2	60.98%	39.02%	100%
Total	78.57%	21.43%	100%

Tabla 6.6 – Poder de discriminación del modelo de regresión

Se puede observar que aproximadamente el 79% de las empresas de la muestra fue clasificada correctamente, mientras que sólo en el 21% de los casos sucedió lo contrario. Nótese también que la probabilidad de clasificar incorrectamente a una empresa del grupo 1 es de un 4.65%, mientras que en el caso del grupo 2 es de un 39%. Esto sugiere que es más difícil reconocer a las empresas que tienen problemas financieros, aunque debemos tomar este resultado con reserva dado que el poder de discriminación se ha evaluado sobre la misma muestra a partir de la cual se desarrolló el modelo.

Este análisis demuestra que a pesar de que los coeficientes de determinación del modelo no tienen valores elevados, su poder de discriminación sigue siendo significativo.



Para finalizar con la validación del modelo se llevaron a cabo las pruebas de significancia de los coeficientes de regresión cuyos resultados se resumen en la Tabla 6.7.

Contraste de Wald	Parámetros Estimados	
	b₁ (F₁)	b₂ (F₂)
Hipótesis Nula	H ₀) $\beta_1 = 0$	H ₀) $\beta_2 = 0$
Estadístico Calculado	$t = -1.9751$	$t = 2.9763$
Valor Crítico	$t_{81,0.975} = 1.9897$	$t_{81,0.95} = 1.6639$
Condición de Rechazo	$ t \geq t_{81,0.975}$	$t \geq t_{81,0.95}$
Resultado	No Rechazar	Rechazar

Contraste de Wilks	Parámetros Estimados	
	b₁ (F₁)	b₂ (F₂)
Hipótesis Nula	H ₀) $\beta_1 = 0$	H ₀) $\beta_2 = 0$
Estadístico Calculado	$\chi^2 = 3.8884$	$\chi^2 = 21.8835$
Valor Crítico	$\chi^2_{1,0.95} = 3.8415$	$\chi^2_{1,0.9} = 2.7055$
Condición de Rechazo	$\chi^2 \geq \chi^2_{1,0.95}$	$\chi^2 \geq \chi^2_{1,0.9}$
Resultado	Rechazar	Rechazar

Tabla 6.7 – Pruebas de significancia de los coeficientes de regresión

En primer lugar, se utilizó el contraste de Wald para el ensayo de la hipótesis nula H_0) $\beta_1 = 0$. En una primera instancia, se podría pensar que en este caso corresponde efectuar un ensayo unilateral dado que se espera que a medida que el valor de la componente F_1 – *Variación Operativa* sea mayor, la probabilidad de que una empresa se presente en concurso preventivo disminuya. De todas maneras, teniendo en cuenta que esta componente tiene una alta correlación con el apalancamiento operativo - el cual es sinónimo de riesgo - y que existen empresas de ambos grupos con altos valores de F_1 , se ha decidido emplear un ensayo bilateral cuya condición de rechazo está dada por la expresión (6.33).

Según se observa en la Tabla 6.7, el valor crítico en este caso sugiere que por una pequeña diferencia no se debe rechazar la hipótesis nula, y por lo tanto, no habría suficientes pruebas estadísticas que demuestren que este componente contribuya en la discriminación de las empresas. De todas formas, antes de tomar una decisión definitiva sobre si debe descartarse esta variable del modelo, más adelante continuaremos con este análisis al evaluar los resultados obtenidos mediante el contraste de Wilks.



El segundo ensayo mediante el contraste de Wald corresponde a la hipótesis nula $H_0) \beta_2 = 0$. A diferencia del ensayo anterior, en este caso hay evidencias suficientes que determinan que se debe efectuar un ensayo unilateral. Está claro que cuanto mayor sea el valor de la componente F_2 – *Variación de la Rentabilidad*, mayor debería ser la probabilidad de que una empresa se presente en concurso preventivo. La condición de rechazo viene dada entonces por la expresión (6.31), y según se observa en la Tabla 6.7, se rechaza la hipótesis nula. Esto último implica que hay suficientes pruebas que determinan que este parámetro no es nulo.

A continuación, veremos los resultados obtenidos mediante el contraste de Wilks. El primer ensayo corresponde a la hipótesis nula $H_0) \beta_1 = 0$. Considerando - como vimos anteriormente - que este ensayo es bilateral, a diferencia del resultado obtenido mediante el contraste de Wald, en este caso se rechaza la hipótesis nula por una pequeña diferencia. Entonces, en principio, existen suficientes pruebas que permiten afirmar que el parámetro en cuestión no es nulo.

Esta situación nos plantea la disyuntiva sobre si debe incluirse la componente F_1 – *Variación Operativa* en el modelo, dado que tanto en el contraste de Wald como en el contraste de Wilks, el valor del estadístico calculado se encuentra muy próximo a su valor crítico siendo los resultados obtenidos opuestos. Para resolver esta cuestión se tuvieron en cuenta dos factores. En primer lugar, debemos considerar que si se adopta para cada ensayo un nivel de significación del 10%, en vez de un 5%, entonces en los dos casos se rechazaría la hipótesis nula y existirían pruebas suficientes para pensar que el parámetro β_1 no es nulo. En segundo lugar, en el caso de que se decida descartar la componente F_1 del modelo, el poder de discriminación disminuye considerablemente como puede observarse en la Tabla 6.8. Por estos motivos, se ha decidido no descartar a esta variable del modelo.

	<i>Clasificaciones Correctas</i>	<i>Clasificaciones Incorrectas</i>	<i>Total</i>
<i>Grupo 1</i>	90.70%	9.30%	100%
<i>Grupo 2</i>	46.51%	53.49%	100%
<i>Total</i>	68.60%	31.40%	100%

Tabla 6.8 – Poder de discriminación del modelo
de regresión al excluir la componente principal F_1

El último ensayo de validación que debemos efectuar mediante el contraste de Wilks corresponde a la hipótesis nula $H_0) \beta_2 = 0$. Como vimos anteriormente, éste es un ensayo unilateral y la condición de rechazo está dada por la expresión (6.36). Según se observa en la Tabla 6.7, la hipótesis nula se rechaza claramente. Por lo tanto, al



igual que en el contraste de Wald, se han encontrado pruebas suficientes que permiten afirmar que este parámetro no es nulo.

En resumen, a partir de esta validación se puede concluir que la solución definitiva sigue siendo aquella representada por la expresión (6.39) y no se encuentran desde el punto de vista estadístico pruebas suficientes para descartar este modelo.

Para finalizar con este análisis, se han determinado los intervalos de confianza de cada parámetro con un nivel de confianza del 95%, siendo los resultados obtenidos:

$$\beta_1 = -0.0265 \pm 0.0267 \quad (6.40)$$

$$\beta_2 = 0.1744 \pm 0.1166 \quad (6.41)$$

Hasta aquí, hemos desarrollado un modelo a partir de las componentes principales que nos permite estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo. Podríamos decir, entonces, que se ha alcanzado el objetivo propuesto y ahora no sólo podemos representar a una empresa para conocer la variación de su operación y de su rentabilidad, sino que también podemos calcular en forma cuantitativa el grado de cercanía de dicha empresa hacia el grupo 1 o el grupo 2. Si bien los coeficientes de determinación indican que el modelo explica un porcentaje reducido de la variable dependiente, ésta no es una razón suficiente para descartar el modelo, sobre todo, teniendo en cuenta que tiene una buena capacidad para discriminar a las empresas.

Ahora bien, si tenemos en cuenta que en realidad cada una de las componentes principales es una combinación lineal de las 45 variables originales, nos podríamos hacer la siguiente pregunta: ¿Qué pasaría si desarrolláramos un modelo de regresión logística únicamente a partir de aquellas variables originales sobre las cuales las componentes cargan significativamente?. El motivo por el cual resulta razonable efectuar esta pregunta, radica en que la varianza de cada componente principal puede verse afectada por los valores que toman algunas variables originalmente consideradas, que por sí solas no permiten clasificar a las empresas en alguno de los dos grupos considerados. No obstante, esto no desmerece el análisis efectuado anteriormente pero se plantea la posibilidad de hallar un nuevo modelo que explique mejor el comportamiento de las empresas de la muestra.

Entonces, con el objetivo de responder esta pregunta, en el próximo apartado analizaremos un nuevo modelo de regresión logística a partir de las variables originales.



6.4.2. Segundo modelo a partir de las variables originales

En el modelo de regresión logística desarrollado a partir de las componentes principales, se obtuvieron coeficientes de determinación que explicaban un porcentaje reducido de la varianza total de las empresas de la muestra. Es posible que esto se deba a la influencia que ejercen sobre las componentes algunas variables que por sí solas no permiten establecer una buena diferenciación entre los dos grupos de empresas considerados.

Para verificar esta hipótesis, analizaremos los resultados que se obtienen al desarrollar un modelo de regresión logística a partir de las variables originales sobre las cuales las componentes principales cargan significativamente. A su vez, más adelante se llevará a cabo un análisis global sobre la totalidad de las variables consideradas con el objetivo de encontrar un modelo que explique mejor la varianza de los datos de la muestra.

En primer lugar, nos centraremos sobre aquellas variables que tienen una correlación mayor o igual a 0.5 con las componentes principales, como se muestra en la Tabla 6.9.

<i>Símbolo</i>	<i>Nombre</i>	<i>Tipo de Variable</i>
X ₁₄	Apalancamiento Operativo	Variable explicativa - Continua
X ₃₃	Variación de la Utilidad Neta	Variable explicativa - Continua
X ₃₉	Variación del Endeudamiento Total	Variable explicativa - Continua
X ₄₁	Variación de la Cobertura de la deuda	Variable explicativa - Continua
X ₄₂	Variación del Margen Operativo	Variable explicativa - Continua
X ₄₃	Variación del Margen Neto	Variable explicativa - Continua
X ₄₅	Variación de la Rentabilidad del Activo	Variable explicativa - Continua
Y	Concurso Preventivo	Variable explicada - Dicotómica

Tabla 6.9 – Primer grupo de variables consideradas en el modelo de regresión

Existen una gran cantidad de combinaciones que pueden efectuarse a partir de estas variables para armar un modelo de regresión. De todas formas, habiendo analizado sólo algunas de ellas, es posible contrastar nuestra hipótesis. Consideremos, por ejemplo, el siguiente modelo que incluye únicamente variables que tienen una correlación media o alta con la componente F_2 – *Variación de la Rentabilidad*:

$$\hat{Y} = \frac{1}{1 + e^{-(0.423 - 0.0581X_{45} + 1.2751X_{39} - 0.2968X_{43})}} \quad (6.42)$$



donde los parámetros han sido estimados por el método de máxima verosimilitud. En primer lugar, se puede verificar que todos los signos de estos parámetros concuerdan con la naturaleza de cada una de las variables consideradas. Por un lado, se verifica que un incremento de la rentabilidad del activo o del margen neto implica una probabilidad menor de que una empresa se presente en concurso preventivo. Por el otro, se puede observar que un incremento del endeudamiento total provoca un aumento de la probabilidad de que una empresa entre en cesación de pagos, como era de esperar.

Habiendo verificado el signo de los parámetros, veremos a continuación algunos aspectos de la validación del modelo. Cabe aclarar, que en este caso el interés del análisis se halla en verificar que con unas pocas variables originales se puede obtener un modelo que tenga un poder de discriminación similar a aquel que se obtuvo a partir de las componentes principales. Por este motivo, la validación del modelo estará dirigida hacia los ensayos de hipótesis globales y hacia las medidas de bondad de ajuste. Por el momento, dejaremos de lado las pruebas de significancia de los coeficientes de regresión para el caso en que analicemos un modelo definitivo.

De esta manera, el primer ensayo corresponde a la hipótesis nula H_0) *El modelo se ajusta correctamente a los datos*. Los estadísticos se calcularon mediante las expresiones (6.13) y (6.15), siendo los resultados los que se muestran en la Tabla 6.10.

<i>Ensayo de hipótesis</i>	<i>H₀) El modelo se ajusta correctamente a los datos</i>	
<i>Estadístico calculado</i>	$\chi^2 = 69.153$	$D = 92.4838$
<i>Valor crítico</i>	$\chi^2_{82;0.95} = 104.1387$	$\chi^2_{82;0.95} = 104.1387$
<i>Condición de Rechazo</i>	$\chi^2 \geq \chi^2_{82;0.95}$	$D \geq \chi^2_{82;0.95}$
<i>Resultado</i>	No Rechazar	No Rechazar

Tabla 6.10 – Resultados del primer ensayo global

En ninguno de los dos casos se rechaza la hipótesis nula, y en principio, podría decirse que no hay suficientes pruebas estadísticas que permitan afirmar que el modelo no se ajusta correctamente a los datos.

El segundo ensayo global del modelo corresponde a la hipótesis nula H_0) $\beta_1 = \beta_2 = \dots = \beta_k = 0$. El estadístico se calculó mediante la expresión (6.17), siendo los resultados de este ensayo los que se muestran en la Tabla 6.11.



Ensayo de hipótesis	$H_0) \beta_1 = \beta_2 = \dots = \beta_k = 0$
Estadístico calculado	$G = 26.7375$
Valor crítico	$\chi^2_{3;0.95} = 7.8147$
Condición de Rechazo	$G \geq \chi^2_{3;0.95}$
Resultado	Rechazar

Tabla 6.11 – Resultados del segundo ensayo global

En este caso, la hipótesis nula se rechaza, y en consecuencia, hay razones para creer que alguno de los parámetros estimados no es nulo.

Finalmente, habiendo efectuado una validación general del modelo, podemos evaluar las medidas de bondad de ajuste. En la Tabla 6.12, se muestran los valores calculados de los distintos coeficientes de determinación.

Medidas de bondad de ajuste	Valor
$R^2_{Clásico}$	0.2660
$R^2_{McFadden}$	0.2243
$R^2_{Aldrich-Nelson}$	0.2372
$R^2_{Cox-Snell}$	0.2672
$R^2_{Nagelkerke}$	0.3563

Tabla 6.12 – Valores obtenidos de los coeficientes de determinación

Resulta interesante observar que el máximo valor del coeficiente de determinación está dado por $R^2_{Nagelkerke} = 0.3563$. Si recordamos que en el modelo desarrollado a partir de las componentes principales, este mismo coeficiente toma un valor de $R^2_{Nagelkerke} = 0.3187$, en principio, podríamos afirmar que la hipótesis planteada al comienzo de este apartado es correcta.

Veamos a continuación el poder de discriminación de este modelo. Adoptando el mismo criterio que en el caso de las componentes principales, en la Tabla 6.13, se resume el porcentaje de clasificaciones correctas e incorrectas efectuadas por el modelo para cada grupo de empresas de la muestra.



	<i>Clasificaciones Correctas</i>	<i>Clasificaciones Incorrectas</i>	<i>Total</i>
Grupo 1	90.70%	9.30%	100%
Grupo 2	60.47%	39.53%	100%
Total	75.58%	24.42%	100%

Tabla 6.13 – Poder de discriminación del modelo de regresión

Se puede observar que un 76% de las empresas han sido clasificadas correctamente, mientras que en un 24% de los casos sucedió lo contrario. Por lo tanto, se refuerza la idea de que éste modelo posee una mayor efectividad para clasificar a las empresas respecto del uso de las componentes principales, dado que con mucha menos información, se logran los resultados similares.

Este razonamiento, nos hace pensar que si bien el análisis de componentes principales fue útil para representar a las empresas y definir sectores que responden a determinadas características, utilizar en forma directa dichos resultados en un análisis de regresión logística puede no ser efectivo. De hecho, hemos comprobado que a partir de unas pocas variables originales se obtiene un poder de discriminación de las empresas muy similar.

En resumen, hemos demostrado que en este estudio en particular resulta difícil combinar los resultados del análisis de componentes principales con el de regresión logística. Esta situación nos plantea entonces un nuevo problema. Puede suceder que existan ciertas variables originales que sean significativas para estimar la probabilidad de que una empresa se presente en concurso preventivo, pero que exhiben una baja ponderación en el análisis de componentes principales.

Para resolver esta cuestión, se ha decidido llevar a cabo un análisis global que contemple a las 45 variables originales. Está claro que teniendo en cuenta todas las combinaciones que se pueden realizar con estas variables para armar un modelo que tenga como máximo 5 variables explicativas, sería muy difícil evaluar y comparar cada uno de ellos mediante el método de prueba y error.

Por tal motivo, se utilizó un criterio con el fin de reducir el campo de búsqueda. Dicho criterio, consiste en calcular para cada una de las variables por separado el coeficiente de determinación clásico correspondiente al modelo de regresión logística con una sola variable explicativa. Con esto, se pretende establecer cuales son las variables que por sí solas pueden explicar un mayor porcentaje de la varianza total de las empresas de la muestra. Es importante destacar que una variable puede mostrar un R^2 bajo al ser considerada en forma aislada, pero cuando se combina con otra variable la información que aporta es mucho mayor. De todas formas, este método sigue siendo útil para orientarnos en la búsqueda, sobre todo, teniendo en cuenta la gran



cantidad de variables consideradas y que muchas de ellas son índices financieros que tienen una alta correlación entre sí.

<i>Variable Explicativa</i>	<i>R²_{Clásico}</i>
X ₂₉ . Rentabilidad del Activo	0.3687
X ₂₇ . Rentabilidad del Capital Empleado (ROCE)	0.3116
X ₂₅ . Rentabilidad del Patrimonio Neto (ROE)	0.3088
X ₄₄ . Variación de la Rentabilidad del Patrimonio Neto	0.2505
X ₂₄ . Margen Neto	0.2386
X ₄₃ . Variación del Margen Neto	0.2326
X ₂₆ . Rentabilidad Comparada	0.2258
X ₁₆ . Endeudamiento Corriente	0.2242
X ₁₅ . Endeudamiento Total	0.2236
X ₂₃ . Margen Operativo	0.1977
X ₄₅ . Variación de la Rentabilidad del Activo	0.1894
X ₁₉ . Apalancamiento Externo	0.1885
X ₃₂ . Variación del Activo	0.1762
X ₂ . Fondo de Maniobra	0.1672
X ₃₃ . Variación de la Utilidad Neta	0.1642
X ₃ . Liquidez Corriente	0.1627
X ₂₀ . Cobertura de la deuda	0.1502
X ₅ . Liquidez Absoluta	0.1464
X ₂₁ . Costo Financiero	0.1110
X ₃₉ . Variación del Endeudamiento Total	0.1037
X ₃₀ . Rentabilidad Operativa del Activo	0.1026
X ₁₇ . Financiación Externa	0.1003
X ₄₀ . Variación de la Financiación Externa	0.0596
X ₁₁ . Rotación de Activos Corrientes	0.0429
X ₈ . Período de Inventario	0.0378
X ₃₅ . Variación de la Inmovilización de Activos	0.0369
X ₃₆ . Variación del Fondo de Maniobra	0.0335
X ₆ . Período Promedio de Cobranza	0.0334
X ₄₂ . Variación del Margen Operativo	0.0332
X ₁₄ . Apalancamiento Operativo	0.0233
X ₁₀ . Rotación de Activos	0.0226
X ₃₈ . Variación de la Rotación de Activos	0.0198
X ₁₈ . Capitalización Externa	0.0127
X ₂₂ . Margen Bruto	0.0102
X ₇ . Período Promedio de Pago	0.0101
X ₁₂ . Rotación de Activos Fijos	0.0093
X ₃₇ . Variación de la Liquidez Corriente	0.0056
X ₁ . Inmovilización de Activos	0.0053
X ₄₁ . Variación de la Cobertura de la deuda	0.0053
X ₃₁ . Variación de las Ventas	0.0038
X ₉ . Ciclo del Efectivo	0.0037
X ₂₈ . Rentabilidad Operativa del Capital Empleado (ROCE)	0.0028
X ₁₃ . Rotación del Patrimonio Neto	0.0008
X ₄ . Liquidez Seca	0.0003
X ₃₄ . Variación del Pasivo	0.0003

Tabla 6.14 – Ranking de los R^2 en las variables originales



En la Tabla 6.14, se muestra el ranking que se obtuvo de los $R^2_{\text{clásico}}$. Lo primero que podemos observar, es que las variables que tienen una alta correlación con las componentes principales se encuentran dispersas, aunque aquellas que están relacionadas con la componente F_2 – *Variación de la Rentabilidad* tienen la mayoría un R^2 mayor que 0.1. Esto se podría explicar teniendo en cuenta que en el modelo de regresión el parámetro estimado de dicha componente es más significativo que el de la componente F_1 – *Variación Operativa*. Nótese además que las primeras variables que por sí solas explican mejor el comportamiento de las empresas, están todas relacionadas con aspectos de la rentabilidad.

A partir de estos resultados, el próximo paso que se realizó fue descartar las últimas 23 variables que presentaban un coeficiente de determinación menor que el resto del conjunto, reduciendo así el análisis a sólo 22 variables. Por último, éstas se agruparon según su correlación para evitar incluir simultáneamente en un modelo variables que aporten el mismo tipo de información. Por ejemplo, no resulta razonable desarrollar un modelo únicamente con variables de liquidez, dado que seguramente estaríamos descartando información valiosa sobre las características operativas, el endeudamiento, la rentabilidad y la evolución de la empresa.

Finalizada esta primera observación de las variables, se llevó a cabo la prueba de diferentes modelos de regresión logística, respetando siempre la premisa de no incluir más de 5 variables explicativas. La comparación de los distintos modelos se realizó, fundamentalmente, según los valores de los coeficientes de determinación y el criterio del autor sobre la importancia y el significado de cada variable.

Luego de haber evaluado un total de 57 modelos posibles (ver Anexo 10.7), se seleccionó uno de ellos como el más significativo. A modo de comentario, resultante interesante destacar que en la mayoría de los modelos en los que se incluía una variable de liquidez, el signo del parámetro estimado de dicha variable no concordaba con su naturaleza. En otras palabras, se encontró que en muchos modelos, una mayor liquidez de las empresas implicaba una mayor probabilidad de que éstas se declaren en cesación de pagos. Éste fue un motivo por el cual se debieron descartar algunos modelos aún cuando presentaban un alto coeficiente de determinación.

De esta manera, en la Tabla 6.15 se muestran las variables originales que conforman el modelo elegido.



<i>Símbolo</i>	<i>Nombre</i>	<i>Tipo de Variable</i>
X ₁₆	Endeudamiento Corriente	Variable explicativa - Continua
X ₂₉	Rentabilidad del Activo	Variable explicativa - Continua
X ₂₁	Costo Financiero	Variable explicativa - Continua
X ₂₃	Margen Operativo	Variable explicativa - Continua
X ₄₄	Variación de la Rentabilidad del Patrimonio Neto	Variable explicativa - Continua
Y	Concurso Preventivo	Variable explicada - Dicotómica

Tabla 6.15 – Variables seleccionadas en el segundo modelo de regresión

<i>Método de Estimación</i>	<i>Parámetro Estimado</i>					
	<i>b₀</i>	<i>b₁(X₁₆)</i>	<i>b₂(X₂₉)</i>	<i>b₃(X₂₁)</i>	<i>b₄(X₂₃)</i>	<i>b₅(X₄₄)</i>
Mínimos Cuadrados	-2.6012	2.4251	-9.0192	13.3503	-3.1242	-0.2422
Mínimos Cuadrados Ponderados	-1.3466	1.4257	-0.2317	7.5672	-1.3062	-0.2490
Máxima Verosimilitud	-2.2748	2.1978	-2.4296	12.3765	-2.7072	-0.2648

Tabla 6.16 – Resultados de la estimación de los parámetros

Los parámetros fueron estimados mediante los tres métodos vistos en el apartado 6.2 Estimación de los parámetros, siendo los resultados los que se muestran en la Tabla 6.16. Nuevamente, se han considerado los parámetros estimados por el método de máxima verosimilitud como representativos del modelo, teniendo en cuenta que en muchas de las pruebas de validación es condición necesaria que la estimación se haya realizado con este método. Por lo tanto, el modelo queda representado mediante la siguiente expresión:

$$\hat{Y} = \frac{1}{1 + e^{-(2.2748 + 2.1978 X_{16} - 2.4296 X_{29} + 12.3765 X_{21} - 2.7072 X_{23} - 0.2648 X_{44})}} \quad (6.43)$$

donde $\hat{Y}$ representa la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo. Por un lado, se observa que un aumento del endeudamiento corriente o del costo financiero, implica una mayor probabilidad de que la empresa se declare en cesación de pagos. Por el otro, un aumento de la rentabilidad del activo, del margen operativo o de la rentabilidad del patrimonio neto, determinan una menor probabilidad de que la empresa tenga problemas financieros. Por lo tanto, en principio, los signos de los parámetros concuerdan con la naturaleza de las variables explicativas.

A continuación, analizaremos los resultados obtenidos en la validación del modelo comenzando con los ensayos de hipótesis globales.



El primer ensayo corresponde a la hipótesis nula H_0) *El modelo se ajusta correctamente a los datos*. Los estadísticos se calcularon mediante las expresiones (6.13) y (6.15), siendo los resultados los que se muestran en la Tabla 6.17.

<i>Ensayo de hipótesis</i>	H_0) <i>El modelo se ajusta correctamente a los datos</i>	
<i>Estadístico calculado</i>	$\chi^2 = 49.0968$	$D = 60.8275$
<i>Valor crítico</i>	$\chi^2_{80,0.95} = 108.6479$	$\chi^2_{80,0.95} = 108.6479$
<i>Condición de Rechazo</i>	$\chi^2 \geq \chi^2_{80,0.95}$	$D \geq \chi^2_{80,0.95}$
<i>Resultado</i>	No Rechazar	No Rechazar

Tabla 6.17 – Resultados del primer ensayo global

En ninguno de los dos casos se rechaza la hipótesis nula, y en principio, podría decirse que no hay suficientes pruebas estadísticas que permitan afirmar que el modelo no se ajusta correctamente a los datos.

<i>Ensayo de hipótesis</i>	H_0) $\beta_1 = \beta_2 = \dots = \beta_k = 0$
<i>Estadístico calculado</i>	$G = 58.3938$
<i>Valor crítico</i>	$\chi^2_{5,0.95} = 11.0705$
<i>Condición de Rechazo</i>	$G \geq \chi^2_{5,0.95}$
<i>Resultado</i>	Rechazar

Tabla 6.18 – Resultados del segundo ensayo global

El segundo ensayo corresponde a la hipótesis nula H_0) $\beta_1 = \beta_2 = \dots = \beta_k = 0$. El estadístico se calculó mediante la expresión (6.17), siendo los resultados de este ensayo los que se muestran en la Tabla 6.18.

En este caso, la hipótesis nula se rechaza claramente, y en consecuencia, hay razones para creer que alguno de los parámetros estimados no es nulo.

Para continuar con la validación del modelo, se calcularon las medidas de bondad de ajuste. En la Tabla 6.19 se muestran los diferentes coeficientes de determinación.



<i>Medidas de bondad de ajuste</i>	<i>Valor</i>
$R^2_{Clásico}$	0.5858
$R^2_{McFadden}$	0.4898
$R^2_{Aldrich-Nelson}$	0.4044
$R^2_{Cox-Snell}$	0.4929
$R^2_{Nagelkerke}$	0.6572

Tabla 6.19 – Valores obtenidos de los coeficientes de determinación

Lo primero que debemos destacar es la ganancia en significación que se ha conseguido con este nuevo modelo, siendo el máximo valor de los coeficientes calculados $R^2_{Nagelkerke} = 0.6572$. Esto significa, que el conjunto de variables seleccionado explica un 65.72% de la varianza total de las empresas de la muestra.

El poder de discriminación del modelo también ha mejorado considerablemente. En la Tabla 6.20 se resumen la cantidad de empresas de la muestra clasificadas correcta e incorrectamente en cada grupo. El criterio utilizado para establecer si una empresa pertenece al grupo 1 o al grupo 2 sigue siendo el mismo. Las empresas que tengan una probabilidad menor o igual a 0.5 de presentarse en concurso preventivo se consideran dentro del grupo 1, mientras que en el caso contrario pertenecen al grupo 2.

	<i>Clasificaciones Correctas</i>	<i>Clasificaciones Incorrectas</i>	<i>Total</i>
Grupo 1	97.67%	2.33%	100%
Grupo 2	76.74%	23.26%	100%
Total	87.21%	12.79%	100%

Tabla 6.20 – Poder de discriminación del modelo de regresión

Se puede observar que un 87% de las empresas han sido clasificadas correctamente, mientras que sólo un 13% se clasificaron erróneamente. A su vez, recordando que este análisis se basa sobre las empresas a partir de las cuales se desarrolló el modelo, la probabilidad de clasificar incorrectamente a las empresas del grupo 2 ha disminuido notablemente, siendo esta última de un 23.26%.

Por último, se efectuaron las pruebas de significancia de los coeficientes de regresión. En la Tabla 6.21 se muestran los resultados de los contrastes de Wald y Wilks.



Contraste de Wald	Parámetros Estimados				
	b ₁ (X ₁₆)	b ₂ (X ₂₉)	b ₃ (X ₂₁)	b ₄ (X ₂₃)	b ₅ (X ₄₄)
Hipótesis Nula	H ₀) $\beta_1 = 0$	H ₀) $\beta_2 = 0$	H ₀) $\beta_3 = 0$	H ₀) $\beta_4 = 0$	H ₀) $\beta_5 = 0$
Estadístico Calculado	$t = 2.2064$	$t = -0.5036$	$t = 1.6817$	$t = -1.0661$	$t = -1.3713$
Valor Crítico	$t_{80,0.95} = 1.6641$	$t_{80,0.95} = 1.6641$	$t_{80,0.95} = 1.6641$	$t_{80,0.95} = 1.6641$	$t_{80,0.95} = 1.6641$
Condición de Rechazo	$t \geq t_{80,0.95}$	$t \leq -t_{80,0.95}$	$t \geq t_{80,0.95}$	$t \leq -t_{80,0.95}$	$t \leq -t_{80,0.95}$
Resultado	Rechazar	No Rechazar	Rechazar	No Rechazar	No Rechazar

Contraste de Wilks	Parámetros Estimados				
	b ₁ (X ₁₆)	b ₂ (X ₂₉)	b ₃ (X ₂₁)	b ₄ (X ₂₃)	b ₅ (X ₄₄)
Hipótesis Nula	H ₀) $\beta_1 = 0$	H ₀) $\beta_2 = 0$	H ₀) $\beta_3 = 0$	H ₀) $\beta_4 = 0$	H ₀) $\beta_5 = 0$
Estadístico Calculado	$\chi^2 = 9.7588$	$\chi^2 = 0.3969$	$\chi^2 = 4.4051$	$\chi^2 = 4.2093$	$\chi^2 = 6.6551$
Valor Crítico	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$
Condición de Rechazo	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$
Resultado	Rechazar	No Rechazar	Rechazar	Rechazar	Rechazar

Tabla 6.21 – Pruebas de significancia de los coeficientes de regresión

Se observa que en el caso de los parámetros estimados correspondientes a las variables X_{16} – Endeudamiento corriente y X_{21} – Costo financiero, en los dos contrastes se rechazan las hipótesis nulas $H_0) \beta_1 = 0$ y $H_0) \beta_3 = 0$.

Por otro lado, los parámetros estimados de las variables X_{23} – Margen operativo y X_{44} – Variación de la rentabilidad del patrimonio neto, han resultado no significativos mediante el contraste de Wald por una pequeña diferencia. Sin embargo, las hipótesis nulas $H_0) \beta_4 = 0$ y $H_0) \beta_5 = 0$, claramente se rechazan en el contraste de Wilks. Teniendo en cuenta que un nivel de significación del 5% puede resultar muy conservador, se calcularon según los estadísticos del contraste de Wald, las probabilidades máximas de rechazar las hipótesis nulas siendo estas verdaderas. Éstas resultaron $\alpha_4^* = 0.1448$ y $\alpha_5^* = 0.0871$, respectivamente. Por lo tanto, desde el punto de vista estadístico existen razones suficientes para pensar que estos parámetros estimados no son nulos.

Nos queda entonces evaluar el parámetro estimado de la variable X_{29} – Rentabilidad del activo. En este caso, tanto en el contraste de Wald como en el de Wilks, no se rechaza la hipótesis nula $H_0) \beta_2 = 0$. De hecho, la probabilidad máxima de rechazar la hipótesis nula cuando esta es verdadera, resulta de $\alpha_2^* = 0.308$ según el estadístico de Wald y de $\alpha_2^* = 0.5287$ según el estadístico de Wilks. Esto sugiere que existen razones para creer que esta variable debería descartarse del modelo, dado que no contribuye en la diferenciación de las empresas. Nótese que hasta aquí, en general, los ensayos efectuados mediante el contraste de Wilks han resultado más contundentes respecto de los contrastes de Wald.



Teniendo en cuenta que los coeficientes de determinación y el poder de discriminación del modelo prácticamente no se ven afectados al descartar la variable X_{29} – *Rentabilidad del activo*, se puede comprobar que efectivamente la información que aporta es despreciable. Entonces, si descartamos esta variable, el modelo definitivo se puede definir mediante la siguiente expresión:

$$\hat{Y} = \frac{1}{1 + e^{-(2.4567 + 2.2813X_{16} + 14.2315X_{21} - 3.563X_{23} - 0.271X_{44})}} \quad (6.44)$$

donde los parámetros han sido estimados por el método de máxima verosimilitud. Nuevamente, los signos de los parámetros estimados se corresponden con la naturaleza de cada una de las variables explicativas consideradas.

Al efectuar los ensayos de hipótesis globales, al igual que en el caso anterior, no se encontraron evidencias suficientes que indiquen que el modelo no se ajusta correctamente a los datos, como tampoco que los parámetros estimados son todos nulos.

En la Tabla 6.22 y la Tabla 6.23, se muestran los coeficientes de determinación y se resume el poder de discriminación del modelo, respectivamente. Como se comentó anteriormente, los resultados obtenidos no varían significativamente al incluir la variable X_{29} – *Rentabilidad del activo*, hecho que determina que su aporte de información a este modelo es despreciable.

<i>Medidas de bondad de ajuste</i>	<i>Valor</i>
$R^2_{Clásico}$	0.5504
$R^2_{McFadden}$	0.4865
$R^2_{Aldrich-Nelson}$	0.4028
$R^2_{Cox-Snell}$	0.4905
$R^2_{Nagelkerke}$	0.6540

Tabla 6.22 – Valores obtenidos de los coeficientes de determinación



	<i>Clasificaciones Correctas</i>	<i>Clasificaciones Incorrectas</i>	<i>Total</i>
Grupo 1	95.35%	4.65%	100%
Grupo 2	74.42%	25.58%	100%
Total	84.88%	15.12%	100%

Tabla 6.23 – Poder de discriminación del modelo de regresión

En la Tabla 6.24 se resumen las pruebas de significancia de los coeficientes de regresión. Se puede observar que en todos los casos se rechaza la hipótesis nula $H_0) \beta_i = 0$, tanto en el contraste de Wald como en el de Wilks, siendo los resultados más significativos que en el modelo anterior.

<i>Contraste de Wald</i>	<i>Parámetros Estimados</i>			
	b₁ (X₁₆)	b₃ (X₂₁)	b₄ (X₂₃)	b₅ (X₄₄)
<i>Hipótesis Nula</i>	$H_0) \beta_1 = 0$	$H_0) \beta_3 = 0$	$H_0) \beta_4 = 0$	$H_0) \beta_5 = 0$
<i>Estadístico Calculado</i>	$t = 2.3395$	$t = 2.1002$	$t = -1.6862$	$t = -1.7514$
<i>Valor Crítico</i>	$t_{81,0.95} = 1.6639$	$t_{81,0.95} = 1.6639$	$t_{81,0.95} = 1.6639$	$t_{81,0.95} = 1.6639$
<i>Condición de Rechazo</i>	$t \geq t_{81,0.95}$	$t \geq t_{81,0.95}$	$t \leq -t_{81,0.95}$	$t \leq -t_{81,0.95}$
<i>Resultado</i>	Rechazar	Rechazar	Rechazar	Rechazar

<i>Contraste de Wilks</i>	<i>Parámetros Estimados</i>			
	b₁ (X₁₆)	b₃ (X₂₁)	b₄ (X₂₃)	b₅ (X₄₄)
<i>Hipótesis Nula</i>	$H_0) \beta_1 = 0$	$H_0) \beta_3 = 0$	$H_0) \beta_4 = 0$	$H_0) \beta_5 = 0$
<i>Estadístico Calculado</i>	$\chi^2 = 9.8813$	$\chi^2 = 6.4451$	$\chi^2 = 6.4641$	$\chi^2 = 15.2369$
<i>Valor Crítico</i>	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$	$\chi^2_{1,0.9} = 2.7055$
<i>Condición de Rechazo</i>	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$	$\chi^2 \geq \chi^2_{1,0.9}$
<i>Resultado</i>	Rechazar	Rechazar	Rechazar	Rechazar

Tabla 6.24 – Pruebas de significancia de los coeficientes de regresión

Para finalizar con el análisis de regresión, se calcularon los intervalos de confianza de cada parámetro con un nivel de confianza del 95%, siendo los resultados obtenidos:

$$\beta_1 = 2.2813 \pm 1.9407 \quad (6.45)$$

$$\beta_3 = 14.2315 \pm 13.4856 \quad (6.46)$$

$$\beta_4 = -3.563 \pm 4.2052 \quad (6.47)$$



$$\beta_5 = -0.271 \pm 0.308 \quad (6.48)$$

En resumen, hemos encontrado finalmente un modelo de regresión logística a partir de las variables originales el cual se define mediante la expresión (6.44). Los resultados obtenidos demuestran una superioridad respecto del modelo desarrollado a partir de las componentes principales, no sólo en la explicación de la varianza total de las empresas de la muestra, sino también en su poder de discriminación.

Se confirma entonces que en este estudio en particular, se pierde una gran cantidad de información al utilizar las componentes principales directamente en el análisis de regresión logística. De todas formas, es importante destacar que a partir de estos dos análisis se han desarrollado dos herramientas de gran utilidad que nos permiten evaluar y resumir las condiciones financieras de una empresa a partir de sus estados contables.



7. PRUEBAS CON UNA NUEVA MUESTRA DE DATOS

El objetivo que se busca en esta sección es evaluar la efectividad en el reconocimiento de las empresas con problemas financieros cuando se aplica el análisis de componentes principales y el de regresión logística en forma conjunta sobre una nueva muestra de datos.

Debido a la dificultad que representa la recopilación de la información, la muestra está compuesta únicamente por 14 empresas de las cuales sólo 3 de ellas se han presentado en concurso preventivo (ver Anexo 10.8 para un detalle de la muestra). Además, debemos considerar que en el caso de dichas empresas los estados contables relevados corresponden a períodos anteriores al 2002. Sin embargo, a pesar de estas limitaciones, el intento de evaluar la efectividad de estos análisis a partir de nuevas observaciones sigue siendo válido, sólo que los resultados que se obtengan deberán interpretarse con precaución.

En primer lugar, se han calculado para cada empresa los valores de las componentes principales F_1 – Variación Operativa y F_2 – Variación de la Rentabilidad, los cuales se han representado en un mismo gráfico como se muestra en la Figura 7.1.

Las tres empresas que se declararon en cesación de pagos se ubican en el sector 2 que corresponde a un elevado apalancamiento operativo acompañado por un deterioro de la rentabilidad. A su vez, la mayoría de las empresas que no tuvieron problemas financieros también se hallan en este sector, y sólo dos de ellas, se encuentran en el sector 6 que responde a una variación operativa positiva y una mejora de la rentabilidad.

Por lo tanto, hasta aquí, bajo el supuesto de que no conocemos a priori el desenlace de la situación particular de cada empresa, podríamos decir que todas aquellas que se encuentran en el sector 2 tienen una mayor predisposición hacia los problemas financieros mientras que ocurre lo contrario con aquellas que se hallan en el sector 6. Sin embargo, para lograr una mayor precisión en la clasificación, deberíamos utilizar el modelo de regresión logística.

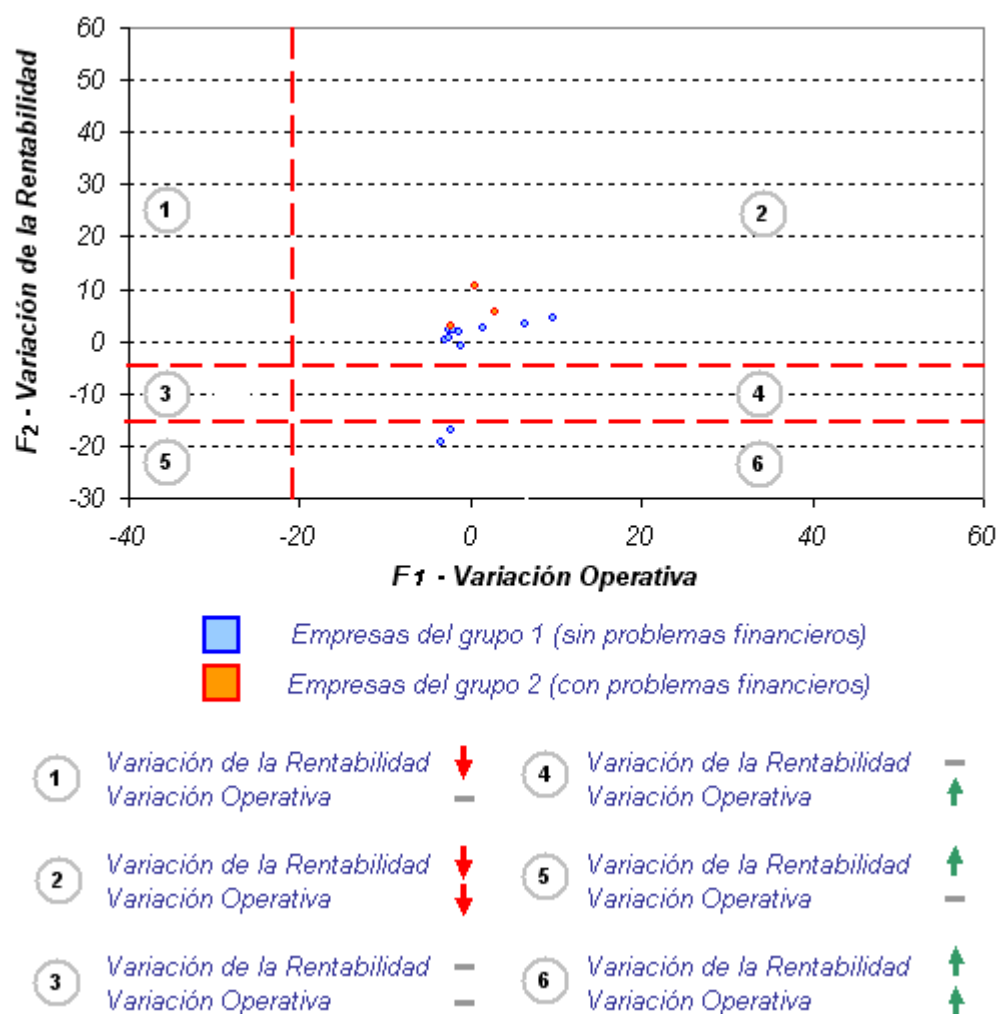


Figura 7.1 – Clasificación del nuevo grupo de empresas

De esta manera, se calculó para cada empresa la probabilidad de que se presente en concurso preventivo en el corto plazo, siendo los resultados aquellos que figuran en la Tabla 7.1.

Por lo tanto, si continuamos adoptando el criterio de que las empresas que tienen una probabilidad de presentarse en concurso preventivo menor o igual a 0.5 se consideran dentro del grupo 1 (sin problemas financieros), y en caso contrario pertenecen al grupo 2 (con problemas financieros), vemos que todas las empresas de la nueva muestra han sido clasificadas correctamente.



<i>Empresa N°</i>	<i>Grupo N°</i>	<i>Probabilidad</i>
1	1	0.0045
2	1	0.2164
3	1	0.1569
4	1	0.3691
5	1	0.2479
6	1	0.0863
7	1	0.2680
8	1	0.4766
9	1	0.1244
10	1	0.3013
11	1	0.2462
12	2	0.5593
13	2	0.7444
14	2	0.5279

Tabla 7.1 – Resultados del modelo de regresión
logística en la nueva muestra

En síntesis, los resultados obtenidos sugieren que las herramientas desarrolladas en secciones anteriores son útiles para reconocer a empresas con problemas financieros más allá de los errores de clasificación esperados en cualquier modelo de regresión. Es importante destacar cómo en este caso en particular los dos análisis se han complementado de manera efectiva, dado que muchas empresas, a pesar de haber sufrido un deterioro de sus rentabilidades, no se habían presentado en concurso preventivo.





8. CONCLUSIONES

A lo largo de este estudio se ha podido comprobar mediante los análisis utilizados que los índices financieros muestran un comportamiento distinto en las empresas con problemas financieros respecto de aquellas que no se presentaron en concurso preventivo, por lo menos hasta un ejercicio previo a dicho suceso. Si bien no todos ellos tienen por sí solos la misma capacidad de discriminar a las empresas, es posible combinar y resumir toda esta información en dos componentes principales que hemos denominado *Variación Operativa* y *Variación de la Rentabilidad*. A través de estas variaciones se puede determinar la evolución que tuvo una empresa en estos aspectos, y así, inferir sobre su predisposición hacia los problemas financieros, aunque no siempre permiten establecer en forma precisa dicha condición.

Los resultados significativos del análisis de componentes principales sugieren que, en principio, no habría razones suficientes como para descartar alguna de las variables originalmente consideradas a la hora de estimar la probabilidad de que una empresa se presente en concurso preventivo. Sin embargo, se logra una mayor significancia si se utilizan unas pocas variables explicativas que en su conjunto tienen un gran poder de discriminación. Esta ganancia de efectividad se debe a que las componentes principales, al ser una combinación lineal de otras 45 variables, contienen información de algunos índices que no contribuyen en la clasificación de las empresas. Por lo tanto, vemos que es fundamental en este tipo de estudios contar con una base de datos completa para evitar introducir limitaciones en el análisis. A su vez, debemos destacar la conveniencia del uso de diferentes herramientas, las cuales, al aplicarse aún sobre la misma muestra, permiten un mejor aprovechamiento de la información.

A lo largo del estudio, se ha podido comprobar que las variables que por sí solas tienen un mayor poder de discriminación son aquellas relacionadas con la rentabilidad. De hecho, la componente principal *Variación de la Rentabilidad* resultó ser más significativa que la *Variación Operativa*. De todas formas, esta información se puede complementar con otro tipo de índices, junto con los cuales se pueden identificar con mayor precisión a las empresas con dificultades financieras.

Luego de haber efectuado un análisis global de todas las variables originales desde el punto de vista de la regresión, se seleccionó un grupo de índices que logran explicar un 65% del comportamiento de las empresas de la muestra. Dicho grupo está compuesto por el *Endeudamiento corriente*, el *Costo financiero*, el *Margen operativo* y la *Variación de la rentabilidad del patrimonio neto*.

Es interesante destacar que, en general, en todos los modelos que se han probado hay una mayor probabilidad de clasificar incorrectamente a una empresa con problemas financieros respecto de aquellas que no los tienen. Seguramente esto pueda explicarse teniendo en cuenta que muchos de los índices presentan una mayor dispersión en las empresas que se han presentado en concurso preventivo. De todas



maneras, el hecho de poder complementar el análisis de componentes principales con el de regresión logística ayuda a reducir la probabilidad de efectuar una clasificación incorrecta. Nótese al respecto que no se han considerado en este estudio los costos vinculados a la toma de decisiones, aunque en el caso de que existan dudas sobre si una empresa presenta problemas financieros, en general, la postura más conservadora indicaría que deberíamos aceptar dicha hipótesis.

El resultado final de este trabajo son dos herramientas basadas en la inferencia estadística a través de las cuales es posible estimar la probabilidad de que una empresa se presente en concurso preventivo en el corto plazo y evaluar las variaciones que ha tenido en su rentabilidad y su operación. Existen diversas aplicaciones para estas herramientas entre las cuales podemos mencionar: (1) el control y seguimiento de la situación financiera de una empresa; (2) complemento para la adjudicación de préstamos; (3) complemento para la toma de decisiones de inversión, fusiones o adquisiciones; (4) complemento en estudios de mercado o sectores económicos; (5) complemento para el análisis financiero de una empresa en un proceso legal de concurso.

Esta investigación se podría expandir si se incluyen en el análisis otras variables explicativas que sean cualitativas, en vez de índices financieros, con el fin de evaluar qué tanto pueden contribuir a una mejor predicción de la cesación de pagos. Otra posibilidad implicaría, como hemos comentado anteriormente, utilizar una herramienta a través de la cual se puedan minimizar los costos vinculados a la toma de una decisión incorrecta. Por último, se podría trasladar este estudio a las empresas de gran tamaño y efectuar comparaciones que permitan establecer cuáles son las características de una estructura financiera más sólida. Como se puede observar, existen diferentes maneras de continuar con este estudio y la estadística ofrece herramientas interesantes para ello.



9. BIBLIOGRAFÍA

- [1] Beaver, William H. 1966. *Financial ratios as predictors of failure*. Empirical Research in Accounting, selected studies (in supplement to the Journal of Accounting Research, January, 1967). Páginas 71-111.
- [2] Altman, Edward I. 1968. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. The Journal of Finance. Volumen 23. Número 4. Páginas 589-609.
- [3] Sydney, Alexander S. 1949. *The effect of size of manufacturing corporation on the distribution of the rate of return*. Review of Economics and Statistics. Volumen 31. Número 3. Páginas 229-235.
- [4] Sturges, H. 1926. *The choice of a class interval*. Journal of the American Statistical Association. Volumen 21. Páginas 65-66.
- [5] Peña, D. 2002. *Análisis de datos multivariantes*. 556 páginas. Editorial McGraw-Hill. ISBN 844-8136-10-1.
- [6] Johnson, D. 2000. *Métodos multivariados aplicados al análisis de datos*. 566 páginas. Editorial Thomson. ISBN 968-7529-90-3.
- [7] García, Roberto M. 2004. *Inferencia estadística y diseño de experimentos*. 734 páginas. Editorial Eudeba. ISBN 950-23-1295-3.
- [8] García, L. 2004. *Las empresas más rentables*. Revista Mercado. Julio '04. Páginas 76-80.
- [9] García, L. 2005. *Las empresas más rentables*. Revista Mercado. Julio '05. Páginas 110-114.
- [10] García, L. 2006. *Las 100 empresas más rentables*. Revista Mercado. Julio '06. Páginas 134-138.
- [11] Hosmer, David W. y Lemeshow, Stanley. 2000. *Applied logistic regression*. 375 páginas. Editorial John Wiley & Sons Inc. Segunda Edición. ISBN 0-471-35632-8.
- [12] Schweser. 2007. *Schweser Study Notes, Level 1, Book 3, Financial Statement Analysis*. Additional material for taking the CFA Exam. 249 páginas. Kaplan Inc. ISBN 1-933677-24-4.



[13] Guzmán, Isidoro. 2004. *Factores explicativos del reparto de dividendos a cuenta en las empresas españolas*. Editado por el Instituto Valenciano de Investigaciones Económicas S.A. 24 páginas. Depósito Legal V-1628-2004. www.ivie.es/downloads/docs/wpasec/wpasec-2004-09.pdf. Página vigente al 19/02/07.

[14] Heine, Max L. 2000. *Predicting financial distress of companies: revisiting the Z-Score and ZETA models*. Stern School of Business, New York University. <http://www.stern.nyu.edu/~ealtman/Zscores.pdf>. Página vigente al 19/02/07.

[15] De la Torre Martínez, J., Navarro, R., Azofra, A., Castro Peña, J. 2002. *Aplicación de los modelos SEPARATE y regresión logística para determinar la relevancia de la información contable en el mercado de capitales: una comparación*. Universidad de Granada. <http://www.ax5.com/antonio/publicaciones/delatorre2002aplicacion.pdf>. Página vigente al 19/02/07.

[16] Kahl, Matthias. 2001. *Financial distress as a selection mechanism: Evidence from the United States*. Paper 16-01. Anderson School of Management, University of California, Los Angeles. <http://repositories.cdlib.org/anderson/fin/16-01>. Página vigente al 19/02/07.



10. ANEXOS

10.1. Composición detallada de la muestra de empresas

En la Tabla 10.1 se muestra un detalle de las empresas relevadas que forman parte de la muestra.

<i>Empresa N°</i>	<i>Grupo N°</i>	<i>Nombre</i>	<i>Período Analizado</i>	<i>Actividad desarrollada</i>
1	1	Alvarez Hnos. S.A.	2005-2004	Molinos y aceites
2	1	Compañía Internacional de Alimentos y Bebidas S.A.	2004-2003	Productos alimenticios
3	1	Establecimiento Metalúrgicos Cavanna S.A.C.I.F.I.	2005-2004	Tecnología y comunicaciones
4	1	Andreani Logística S.A.	2004-2003	Transporte
5	1	Compañía de Servicios Telefónicos S.A.	2005-2004	Telecomunicaciones
6	1	Compumundo S.A.	2005-2004	Comercio minorista
7	1	Caputo S.A.	2005-2004	Construcciones
8	1	Ediar S.A.	2005-2004	Imprentas y editoriales
9	1	Agrometal S.A.I.	2005-2004	Máquinas y equipos (bienes de capital)
10	1	Electromac S.A.	2005-2004	Máquinas y equipos (bienes de capital)
11	1	Gijon S.A.	2005-2004	Construcciones
12	1	Green S.A.	2005-2004	Construcciones
13	1	Esat S.A.	2004-2003	Productos industrias plástica y química
14	1	Grafex S.A.	2004-2003	Imprentas y editoriales
15	1	Lihue Ingeniería S.A.	2005-2004	Máquinas y equipos (bienes de capital)
16	1	Laboratorio LKM S.A.	2004-2003	Laboratorios
17	1	Guilford Argentina S.A.	2005-2004	Textil y calzado
18	1	Rovella Carranza S.A.	2005-2004	Construcciones
19	1	Yar Construcciones S.A.	2005-2004	Construcciones
20	1	Mardi S.A.	2004-2003	Industria pesquera
21	1	Mercoplast S.A.	2005-2004	Productos industrias plástica y química
22	1	Bonafide Golosinas S.A.	2005-2004	Productos alimenticios
23	1	Bonesi S.A.	2005-2004	Artículos del hogar (líneas blanca y marrón)
24	1	Molinos Juan Semino S.A.	2004-2003	Molinos y aceites
25	1	City Pharma S.A.	2005-2004	Comercio minorista
26	1	Morixe Hnos. S.A.	2005-2004	Molinos y aceites
27	1	Coniglio S.A.	2005-2004	Textil y calzado
28	1	Curtiduría A. Gaita S.R.L.	2005-2004	Curtiembres y artículos de cuero
29	1	Domec S.A.I.C. y F.	2005-2004	Artículos del hogar (líneas blanca y marrón)
30	1	Dulcor S.A.	2005-2004	Productos alimenticios
31	1	Distribuidora Santa Bárbara S.A.	2005-2004	Industria pesquera
32	1	Outdoors S.A.	2004-2003	Textil y calzado
33	1	Frutucumán S.A.	2003-2002	Exportadoras e importadoras
34	1	García Reguera S.A.	2005-2004	Comercio mayorista
35	1	Instituto Rosenbusch S.A.	2005-2004	Servicios de la salud
36	1	Insumos Agroquímicos S.A.	2005-2004	Comercio minorista
37	1	Industria Textil Argentina (INTA) S.A.	2005-2004	Textil y calzado
38	1	SAT Médica S.A.	2005-2004	Servicios de la salud
39	1	Leyden S.A.I.C. y F.	2005-2004	Máquinas y equipos (bienes de capital)
40	1	Lodge S.A.	2004-2003	Agrícolas y ganaderas
41	1	Longvie S.A.	2005-2004	Artículos del hogar (líneas blanca y marrón)
42	1	Ovoprot International S.A.	2004-2003	Productos alimenticios
43	1	Magalcuer S.A.	2005-2004	Curtiembres y artículos de cuero
44	2	Aero Vip S.A.	2003-2002	Transporte
45	2	Alunamar S.A.	2005-2004	Industria pesquera
46	2	American Falcon S.A.	2003-2002	Transporte
47	2	AS Sistemas S.A.	2003-2002	Tecnología y comunicaciones
48	2	Bascoy S.A.	2003-2002	Transporte
49	2	Cartex S.A.	2004-2003	Textil y calzado
50	2	Casamen S.A.	2003-2002	Productos alimenticios
51	2	Celeritas S.A.	2004-2003	Servicios de la salud
52	2	Comercial Mendoza S.A.	2003-2002	Artículos del hogar (líneas blanca y marrón)
53	2	Crédito José C. Paz S.A.	2003-2002	Construcciones
54	2	D'Vigi S.A.	2004-2003	Comercio minorista
55	2	Drogueria Sigma S.A.	2003-2002	Comercio minorista
56	2	Ecourban S.A.	2004-2003	Residuos
57	2	El Manzanar de Macedo S.A.	2004-2003	Productos alimenticios
58	2	Espejos Versailles S.A.	2003-2002	Vidrio y mat. construcción
59	2	FrigoFruit S.A.	2003-2002	Agrícolas y ganaderas
60	2	Humberto Nicolás Fontana S.A.C.	2004-2003	Artículos del hogar (líneas blanca y marrón)

Tabla 10.1 – Empresas relevadas que forman parte de la muestra (continúa en la página siguiente)



<i>Empresa N°</i>	<i>Grupo N°</i>	<i>Nombre</i>	<i>Período Analizado</i>	<i>Actividad desarrollada</i>
61	2	Impresiones Arco Iris Córdoba S.A.	2003-2002	Imprentas y editoriales
62	2	Industrias Badar S.A.	2003-2002	Tecnología y comunicaciones
63	2	Diabolo Menthe S.R.L.	2003-2002	Textil y calzado
64	2	La Tribu S.R.L.	2003-2002	Productos alimenticios
65	2	Loucen International S.A.	2004-2003	Bebidas y jugos
66	2	Luicar S.R.L.	2003-2002	Hotelería y turismo
67	2	Manfisa Mandataria y Financiera S.A.	2003-2002	Construcciones
68	2	Norte Asistencia Empresaria S.A.	2003-2002	Correo
69	2	Parmalat Argentina S.A.	2003-2002	Lácteos
70	2	Pto. S.A.	2004-2003	Residuos
71	2	Redes Excon S.A.	2003-2002	Distribución y transporte de gas
72	2	Sanatorio Ezeiza S.A.	2004-2003	Servicios de la salud
73	2	Sanatorio Modelo Quilmes S.A.	2004-2003	Servicios de la salud
74	2	Security Consulting S.A.	2003-2002	Tecnología y comunicaciones
75	2	Sepia Beauty S.A.	2004-2003	Limpieza y cosméticos
76	2	Sol de Brasa S.A.	2005-2004	Agrícolas y ganaderas
77	2	Sycon Argentina S.A.	2003-2002	Distribución y transporte de gas
78	2	UOL Sinectis S.A.	2004-2003	Tecnología y comunicaciones
79	2	Yearling S.A.	2003-2002	Servicios de seguridad
80	2	Fundición de Aceros S.A.	2003-2002	Metalúrgicas y siderúrgicas
81	2	Inmar S.A.	2003-2002	Construcciones
82	2	Carpintería Metálica San Eduardo S.A.	2003-2002	Vidrio y mat. construcción
83	2	Marmolería Sierra Chica S.A.	2003-2002	Minero
84	2	Avaca S.A.	2003-2002	Textil y calzado
85	2	Bellas S.A.	2003-2002	Textil y calzado
86	2	Ianson S.A.	2004-2003	Textil y calzado

Tabla 10.1 – Empresas relevadas que forman parte de la muestra (cont.)

10.2. Descripción de los índices financieros considerados

10.2.1. Índices de liquidez

Mediante este grupo de variables se pretende evaluar la capacidad que posee la empresa para satisfacer sus obligaciones a corto plazo.

X_1 – *Inmovilización de activos*. Es un indicador de la proporción de la inversión en activos fijos respecto del total de activos. Un valor muy alto de este índice podría implicar que la empresa tiene poca capacidad para afrontar sus obligaciones de corto plazo.

X_2 – *Fondo de maniobra*. Es un indicador de la proporción de la inversión neta en capital de trabajo respecto del total de activos. Un valor negativo y cercano a la unidad de este índice podría interpretarse como un signo de problemas de liquidez ya sea: (1) porque la empresa tiene poco activos corrientes para operar; (2) porque no puede pagar sus deudas a corto plazo lo cual se traduce en un aumento considerable del pasivo corriente; (3) una combinación de (1) y (2).

X_3 – *Liquidez corriente*. Mide la relación entre los activos corrientes y los pasivos corrientes. Un valor muy bajo podría significar que la empresa posee dificultades para afrontar sus obligaciones de corto plazo.



X_4 – *Liquidez seca*. Se mantiene la misma relación que en el caso anterior pero se excluyen los bienes de cambio de los activos corrientes. El objetivo es contar con una medida más exigente a la hora de evaluar la capacidad de la empresa para afrontar sus obligaciones de corto plazo. Nuevamente, un valor muy bajo de este índice podría implicar que la empresa posee problemas de liquidez.

X_5 – *Liquidez absoluta*. Mide la relación entre las disponibilidades y los pasivos corrientes de la empresa. En este caso sólo se considera el rubro más líquido del activo para evaluar la capacidad de afrontar las obligaciones de corto plazo. Un valor muy bajo estaría indicando que la empresa no cuenta con suficiente dinero de caja y bancos como para cubrir sus deudas de corto plazo.

10.2.2. Índices operativos

El objetivo de estas variables es evaluar la liquidez operativa de la empresa y medir la eficiencia de los recursos utilizados para generar ingresos.

X_6 – *Período promedio de cobranza*. Mide la relación entre los créditos por ventas y las ventas de la empresa. Es un indicador del tiempo promedio que tardan los créditos por venta en convertirse en efectivo. Un valor muy alto de este índice puede implicar que la empresa tiene dificultades para cobrar los créditos otorgados en sus ventas.

X_7 – *Período promedio de pago*. Mide la relación entre las deudas comerciales y las compras de la empresa. En aquellos casos donde las compras de materia prima no aplican, como por ejemplo en empresas de servicios, éstas se reemplazan por el costo de ventas. Es un indicador del tiempo promedio que demora la empresa en pagar sus deudas comerciales y al mismo tiempo muestra que tanto utiliza a sus proveedores como fuente de financiamiento. Un valor muy alto de este índice podría implicar que la empresa tiene un alto riesgo en el cumplimiento de sus compromisos con los proveedores.

X_8 – *Período de inventario*. Mide la relación entre el inventario promedio que mantiene la empresa y el costo de ventas. Es un indicador del tiempo promedio que tarda el inventario en transformarse en crédito. Un valor muy alto de este índice puede implicar que la empresa tiene problemas para vender sus productos en el mercado.

X_9 – *Ciclo del efectivo*. Se define como la suma del período promedio de cobranza y los días de inventario, menos el período promedio de pago. Es una medida del tiempo promedio que transcurre desde que la empresa le paga a sus proveedores hasta que les cobra a sus clientes, es decir, representa el tiempo promedio que demora en recuperarse el efectivo utilizado en las operaciones y durante cual se deben financiar los bienes de cambio y los créditos. Valores positivos y muy grandes de esta variable podrían implicar: (1) que la empresa tiene un nivel de inventario elevado; (2) que tarda mucho tiempo en cobrarle a sus clientes; (3) una combinación de (1) y (2). A su vez,



valores negativos y muy grandes de esta variable podrían implicar dificultades para cumplir el pago con sus proveedores, hecho que se traduce en un período promedio de pago muy alto.

X_{10} – *Rotación de activos*. Mide la relación entre las ventas y los activos de la empresa. Es un indicador de la eficiencia que tiene la empresa para generar ingresos, es decir, una medida de las ventas que generan los activos. Un valor muy bajo de este índice podría implicar que la empresa posee dificultades para generar ingresos ya que ha efectuado una gran inversión en activos que en su conjunto no generan las ventas esperadas.

X_{11} – *Rotación de activos corrientes*. Mide la relación entre las ventas y los activos corrientes de la empresa. Es de esperar que los valores de este índice se encuentren cercanos al promedio del sector. Un valor muy alto del mismo podría implicar que la empresa se ha desprendido de gran parte de sus activos corrientes como consecuencia de problemas de liquidez. Por otro lado, un valor muy bajo podría ser consecuencia de una baja considerable en las ventas del período.

X_{12} – *Rotación de activos fijos*. Mide la relación entre los activos fijos y las ventas de la empresa. Es un indicador de la inmovilización de capital que tiene la empresa. Es de esperar que las empresas que no muestren problemas financieros tengan un valor de este índice similar y cercano al promedio del sector. Un valor muy alto del mismo podría implicar que la empresa mantiene demasiado capital inmovilizado en sus activos. A su vez, valores muy bajos pueden implicar que la empresa utiliza equipos obsoletos o que deberá efectuar inversiones en el corto plazo para aumentar la capacidad productiva y soportar el crecimiento de las ventas.

X_{13} – *Rotación del patrimonio neto*. Mide la relación entre el patrimonio neto y las ventas de la empresa. Es un indicador de la magnitud de los ingresos generados respecto de la inversión efectuada en el negocio mediante aportes de los socios y fondos autogenerados. Es de esperar que las empresas tengan un valor de este índice cercano al promedio del sector. Un valor muy alto podría implicar que la empresa no genera suficientes ingresos en comparación con la inversión propia efectuada en el negocio. Por el contrario, un valor muy bajo podría implicar que la empresa ha consumido la mayor parte de su patrimonio neto a causa de sucesivos períodos con pérdidas y que se ha endeudado para mantener sus operaciones. A su vez, un valor negativo representa un caso extremo de la situación anterior en la cual efectivamente se ha consumido todo el patrimonio neto y la empresa se financia únicamente con deuda.



X_{14} – *Apalancamiento operativo*. Mide la relación en valor absoluto entre la variación de la utilidad operativa y la variación de las ventas de la empresa⁷. Es un indicador de como se ven afectadas las utilidades operativas por un cambio en las ventas, es decir, cuantifica de manera aproximada el riesgo del negocio entendiendo como tal a la variabilidad de los ingresos operativos. En cierta forma también refleja la proporción de costos fijos respecto de los costos variables que utiliza la empresa. Por ejemplo, si una variación en las ventas produce exactamente la misma variación en las utilidades, es decir que el índice tiene un valor unitario, estaríamos frente a un caso en donde la empresa opera únicamente con costos variables. Un valor muy alto de apalancamiento operativo implica mayor riesgo de negocio dado que la dispersión de la utilidad operativa es mayor⁸.

10.2.3. Índices de endeudamiento

El objetivo de este conjunto de variables es evaluar la estructura de financiamiento que posee la empresa, y junto con ella, la capacidad de cumplir con sus obligaciones a largo plazo.

X_{15} – *Endeudamiento total*. Mide la relación entre el pasivo y el activo de la empresa. Es un indicador de la estructura de financiamiento, es decir, que proporción de la inversión total en activos ha sido financiada con deuda. Un valor muy alto de este índice determina un mayor riesgo en el cumplimiento de sus obligaciones.

X_{16} – *Endeudamiento corriente*. Mide la relación entre el pasivo corriente y el activo de la empresa. Es un indicador del porcentaje de los activos que son financiados con deuda a corto plazo. Un valor muy alto de este índice podría implicar que la empresa no es efectiva en la elección de las fuentes de financiamiento dado que estaría financiando activos fijos con deuda a corto plazo.

X_{17} – *Financiación externa*. Mide la relación entre el pasivo y las ventas de la empresa. Es un indicador de la cantidad de deuda existente por cada peso generado en las ventas. Un valor muy alto de este índice podría implicar que la empresa tiene dificultades para cumplir con sus obligaciones.

X_{18} – *Capitalización externa*. Mide la relación entre el pasivo no corriente y la suma de éste con el patrimonio neto de la empresa. Es un indicador de la proporción de deuda a largo plazo presente en los recursos permanentes de la empresa. Un valor muy alto de este índice podría implicar que la empresa ha agotado gran parte o la totalidad de sus recursos propios y que prácticamente se financia sólo con deuda, hecho que aumenta el riesgo en el cumplimiento de sus obligaciones.

⁷ Generalmente se considera un promedio de las variaciones que ha mostrado la empresa entre dos ejercicios consecutivos en un período mayor o igual a 5 años. En nuestro caso, debido a la falta de información, se tuvieron en cuenta las variaciones entre los dos ejercicios económicos a partir de los cuales se relevaron el resto de los índices.

⁸ Nótese que el riesgo del negocio no incluye efectos de financiamiento.



X_{19} – *Apalancamiento externo*. Mide la relación entre el patrimonio neto y el pasivo de la empresa. Es una variable similar a la de endeudamiento total que indica la preferencia de la empresa al uso de deuda como fuente de financiamiento (valores menores a la unidad) respecto del uso de fuentes de capital propio (valores mayores a la unidad). Un valor muy bajo de este índice podría implicar un mayor riesgo en el cumplimiento de las obligaciones de la empresa. A su vez, un valor negativo representa un caso extremo de la situación anterior en la cual la empresa ha consumido todo su patrimonio neto y por lo tanto se financia únicamente con deuda aumentando considerablemente el riesgo de la inversión efectuada por terceros.

10.2.4. Índices de cobertura

El objetivo de estas variables es evaluar la capacidad de la empresa para satisfacer los pagos de su deuda y los cargos financieros que ésta genera.

X_{20} – *Cobertura de la deuda*. Mide la relación entre la utilidad operativa y el pasivo de la empresa. Es un indicador de la capacidad que posee la empresa para cumplir con sus obligaciones a través de la ganancia obtenida en el desarrollo de su actividad. Un valor muy bajo o negativo de este índice podría implicar que la empresa no genera suficientes fondos operativos para cubrir su deuda, y por ende, tiene un mayor riesgo en el cumplimiento de sus pagos futuros.

X_{21} – *Costo financiero*. Mide la relación entre los intereses anuales y el pasivo de la empresa. Es un indicador del costo aproximado que afronta la empresa al financiarse con fondos de terceros. Un valor muy alto de este índice podría implicar: (1) que la empresa no es efectiva en la elección de sus fuentes de financiamiento; (2) un mayor riesgo en el pago de los cargos financieros; (3) una combinación de (1) y (2).

10.2.5. Índices de rentabilidad

Mediante estas variables se pretende evaluar la capacidad que posee la empresa para generar ganancias.

X_{22} – *Margen bruto*. Mide la relación entre la utilidad bruta y las ventas de la empresa. Es un indicador de la proporción del costo de la mercadería vendida o de los servicios prestados, según sea el caso, respecto de las ventas generadas. Un valor muy bajo de este índice podría implicar que la empresa posee dificultades para generar ingresos brutos, aunque no deben descartarse las características propias del negocio.

X_{23} – *Margen operativo*. Mide la relación entre la utilidad operativa y las ventas de la empresa. Es un indicador de la eficiencia operativa que posee la empresa para generar ganancias, es decir, que proporción de las ventas representan ingresos luego de deducir el costo de venta, los gastos administrativos, los gastos comerciales y otros



gastos operativos. Un valor muy bajo o negativo de este índice podría implicar que la empresa no es eficiente para generar ganancias debido a la magnitud de sus gastos operativos, aunque nuevamente hay que contemplar el tipo de actividad que ésta desarrolla.

X_{24} – *Margen neto*. Mide la relación entre la utilidad neta y las ventas de la empresa. Es un indicador de la proporción de las ventas que representan ingresos netos para la empresa luego de sumarle o restarle a la utilidad operativa, según corresponda, los resultados financieros y otros ingresos y egresos que no surjan de la operación. Un valor muy bajo o negativo de este índice podría implicar que la empresa tiene dificultades para generar fondos cuyo origen puede ser una ineficiencia operativa, un alto costo financiero, cuestiones excepcionales ajenas a la operación del negocio, o bien, una combinación de todos estos.

X_{25} – *Rentabilidad del patrimonio neto*. Mide la relación entre la utilidad neta y el patrimonio neto de la empresa. Es un indicador de la magnitud de la utilidad neta generada respecto de la inversión efectuada en el negocio mediante aportes de los socios y fondos autogenerados. Es de esperar que los valores de este índice se encuentren cercanos al promedio del sector. Un valor muy alto podría implicar que la empresa se encuentra en un proceso de recuperación luego de haber atravesado sucesivos períodos con pérdida que derivaron en una reducción considerable de su patrimonio neto. Un valor muy bajo de este índice podría implicar que la empresa no genera suficientes utilidades en comparación con la inversión efectuada con fondos propios. A su vez, un valor negativo es un caso extremo de la situación anterior en la cual la empresa en el desarrollo de su actividad ha consumido una parte de los fondos propios.

X_{26} – *Rentabilidad comparada*. Mide la diferencia porcentual entre la rentabilidad del patrimonio neto de la empresa y la rentabilidad promedio del patrimonio neto de otras empresas representativas del mercado que desarrollan la misma actividad. Con esta variable se pretende evaluar el desempeño de la empresa respecto del promedio del sector. Un valor muy alto de este índice podría implicar que la empresa tiene una mejor performance que el resto. Por otro lado, un valor muy bajo podría indicar que la empresa tiene un desempeño inferior al promedio del sector.

X_{27} – *Rentabilidad del capital empleado*. Mide la relación entre la utilidad neta y la suma del patrimonio neto con el pasivo no corriente de la empresa. Es un indicador de la magnitud de la utilidad neta generada respecto del capital empleado en la actividad de la empresa, considerando dicho capital como los aportes de los socios, los fondos autogenerados y aquellos que fueron aportados por terceros mediante deuda a largo plazo. Un valor muy bajo de este índice podría implicar que la empresa no genera suficientes utilidades en comparación con el capital empleado. Un valor negativo refleja que la empresa ha consumido una parte del capital empleado en el desarrollo de su actividad.



X_{28} – *Rentabilidad operativa del capital empleado*. Mide la relación entre la utilidad operativa y la suma del patrimonio neto con el pasivo no corriente de la empresa. Es un índice similar al anterior pero en este caso se evalúa la magnitud de la utilidad operativa generada respecto del capital empleado en la empresa, y por lo tanto, no se considera el tipo de financiamiento elegido ni los resultados extraordinarios. Un valor muy bajo de este índice podría implicar que la empresa no genera suficientes utilidades en comparación con el capital empleado. Un valor negativo refleja que la empresa ha consumido una parte del capital empleado en el desarrollo de su actividad.

X_{29} – *Rentabilidad del activo*. Mide la relación entre la utilidad neta y los activos de la empresa. Es un indicador de la capacidad que tiene la empresa en su conjunto de generar utilidades, es decir, cuantifica el retorno de la inversión total. Un valor muy bajo de este índice podría implicar que la empresa no es eficiente para generar utilidades. Un valor negativo representa una situación en la cual la empresa ha consumido fondos en el desarrollo de su actividad.

X_{30} – *Rentabilidad operativa del activo*. Mide la relación entre la utilidad operativa y los activos de la empresa. Es un índice similar al anterior pero en este caso, al considerar la utilidad operativa, no se tienen en cuenta el costo de financiamiento u otros resultados extraordinarios de la empresa para medir su capacidad de generar utilidades. El objetivo es contar con una medida de la rentabilidad que sea independiente del tipo de financiamiento elegido. Nuevamente, un valor muy bajo de este índice podría implicar que la empresa no es eficiente para generar utilidades y un valor negativo representa una situación en la cual se han consumido fondos.

10.2.6. Índices de evolución

El objetivo de estas variables es medir el crecimiento de la empresa y la evolución de los índices financieros más representativos entre los dos ejercicios económicos relevados.

X_{31} – *Variación de las ventas*. Es un indicador del crecimiento de la empresa. Un valor muy alto y negativo de este índice demuestra que la empresa tuvo dificultades para mantener su nivel de ingresos.

X_{32} – *Variación del activo*. Es un indicador del crecimiento de la empresa. Un valor muy alto y negativo de este índice podría implicar que la empresa tiene dificultades para generar fondos.

X_{33} – *Variación de la utilidad neta*. Es un indicador del crecimiento de la empresa. Un valor muy alto y negativo de este índice demuestra que la empresa tuvo dificultades para mantener su nivel de utilidades.



X_{34} – *Variación del pasivo*. Es un indicador del incremento de la deuda de la empresa. Un valor muy alto de este índice podría implicar un mayor riesgo en el cumplimiento de sus obligaciones, aunque debe considerarse que el aumento de la deuda puede formar parte del crecimiento natural de la empresa y/o de un cambio estratégico en su estructura de financiamiento.

X_{35} – *Variación de la inmovilización de activos*. Un valor muy alto de este índice podría implicar un deterioro en la liquidez de la empresa para responder a sus obligaciones de corto plazo.

X_{36} – *Variación del fondo de maniobra*. Un valor muy alto y negativo de este índice podría implicar un deterioro en la liquidez de la empresa ya sea: (1) porque debe invertir en capital operativo sin generar mayores ingresos; (2) porque sus pasivos corrientes han aumentado considerablemente; (3) una combinación de (1) y (2).

X_{37} – *Variación de la liquidez corriente*. Un valor muy alto y negativo de este índice podría implicar un deterioro en la liquidez de la empresa para cumplir con sus obligaciones de corto plazo.

X_{38} – *Variación de la rotación de activos*. Un valor muy alto y negativo de este índice podría implicar un deterioro en la eficiencia de los activos para generar ingresos.

X_{39} – *Variación del endeudamiento total*. Un valor muy alto de este índice determina un cambio en la estructura de financiamiento de la empresa mediante el incremento de su deuda, hecho que se traduce en un mayor riesgo en el cumplimiento de sus pagos.

X_{40} – *Variación de la financiación externa*. Un valor muy alto de este índice significa que la empresa ha incrementado su deuda para generar ingresos hecho que podría implicar un aumento del riesgo en el cumplimiento de sus obligaciones.

X_{41} – *Variación de la cobertura de la deuda*. Un valor muy alto y negativo de este índice podría implicar que la empresa no genera suficientes fondos operativos para cubrir su deuda lo que determina un mayor riesgo en el cumplimiento de sus pagos futuros.

X_{42} – *Variación del margen operativo*. Un valor muy alto y negativo de este índice podría ser consecuencia de un aumento considerable de los costos operativos, hecho que en última instancia podría interpretarse como un deterioro en la eficiencia operativa de la empresa para generar utilidades.

X_{43} – *Variación del margen neto*. Un valor muy alto y negativo de este índice podría implicar que la empresa tuvo dificultades para generar utilidades cuyo origen puede



ser una ineficiencia operativa, un alto costo financiero, cuestiones excepcionales ajenas a la operación del negocio, o bien, una combinación de todos estos.

X_{44} – *Variación de la rentabilidad del patrimonio neto*. Un valor muy alto y negativo de este índice podría implicar que la empresa tuvo dificultades para generar utilidades dado que la magnitud de la utilidad neta disminuyó considerablemente respecto de la inversión efectuada en la empresa mediante aportes de los socios y fondos autogenerados.

X_{45} – *Variación de la rentabilidad del activo*. Un valor muy alto y negativo de este índice podría implicar un deterioro en la capacidad de la empresa para generar utilidades.

10.2.7. Otros índices

En este apartado se mencionan aquellas variables frecuentes en la bibliografía que fueron inicialmente consideradas para formar parte del estudio pero que luego se decidió no incluirlas en el análisis por diferentes motivos que se explican a continuación.

Rotación de inventario. Mide la relación entre el costo de ventas y el inventario promedio. Éste índice es la inversa del Período de Inventario y se ha descartado del análisis porque no aporta información nueva para discriminar a empresas con problemas financieros.

Rotación de deudas comerciales. Mide la relación entre las compras y las deudas comerciales de la empresa. Éste índice es la inversa del Período Promedio de Pago y se ha descartado del análisis porque no aporta información nueva para discriminar a empresas con problemas financieros.

Rotación de créditos. Mide la relación entre las ventas y los créditos por ventas de la empresa. Éste índice es la inversa del Período Promedio de Cobranza y se ha descartado del análisis porque no aporta información nueva para discriminar a empresas con problemas financieros.

Solvencia. Mide la relación entre el patrimonio neto y los activos de la empresa. Es un indicador de la proporción de la inversión total en activos que ha sido financiada con capital propio de la empresa mediante aportes de los socios y fondos autogenerados. Si bien es un índice bastante común en la bibliografía, se ha descartado del análisis dado que es complementario al índice de endeudamiento total (la suma de ambos es igual a la unidad) y a los efectos de discriminar a empresas con problemas financieros no aporta información nueva.



Cobertura de intereses. Mide la relación entre la utilidad operativa y los intereses de la empresa. Es un indicador de la capacidad que tiene la empresa de pagar los intereses de sus deudas a través de las utilidades que genera. Éste índice se ha descartado del análisis dado que no permite diferenciar en forma unívoca a aquellas empresas que no tienen intereses que pagar respecto de aquellas que no han generado utilidades para cubrirlos. Por esta razón, la variable en cuestión no permite discriminar a empresas con problemas financieros. De todas maneras, al considerar el índice de Costo Financiero, se puede obtener una magnitud del riesgo involucrado en la cobertura de los intereses.

Cobertura de la deuda a largo plazo. Mide la relación entre la utilidad operativa y el pasivo no corriente de la empresa. Es un indicador de la capacidad que tiene la empresa de cumplir con sus obligaciones de largo plazo a través de las utilidades que genera. Éste índice se ha descartado del análisis dado que no permite diferenciar en forma unívoca a aquellas empresas que no tienen deuda a largo plazo respecto de aquellas que no han generado utilidades para cumplir con los pagos de ésta. Por este motivo, la variable en cuestión no permite discriminar a empresas con problemas financieros. De todas formas, al considerar el índice de Cobertura de la Deuda se obtiene una medida similar a la que proporciona esta variable.

Productividad. Mide la relación entre los gastos operativos y las ventas de la empresa. Es un indicador de la eficiencia operativa que tiene la empresa para generar ingresos. Éste índice se ha descartado dado que es muy similar al de margen operativo y a los efectos de discriminar a empresas con problemas financieros no aporta información nueva.

10.3. Rentabilidades del Patrimonio Neto por tipo de actividad

En la Tabla 10.2 se muestra un detalle de las rentabilidades del patrimonio neto (*ROE*) promedio de algunas empresas representativas del mercado según el tipo de actividad económica que desarrollan. Éstas se han utilizado para calcular el índice financiero denominado *Rentabilidad Comparada* para cada una de las empresas de la muestra. La información de esta tabla se construyó sobre la base de tres publicaciones de la revista *Mercado* [8], [9] y [10].



Actividad desarrollada	Año		
	2003	2004	2005
Agrícolas y ganaderas	26.81	29.23	27.02
Artículos del hogar (líneas blanca y marrón)	68.18	29.17	39.83
Automotores y autopartes	31.09	28.27	29.64
Bebidas y jugos	33.51	34.01	29.01
Celulosa y papel	34.75	17.28	18.87 (*)
Comercio mayorista	38.81	36.43	33.44
Comercio minorista	51.69	32.95	67.25
Concesionario vial	86.56 (*)	95.16	103.89 (*)
Construcciones	94.78	47.44	649.63
Correo	44.75	48.79 (*)	53.27 (*)
Curtiembres y artículos de cuero	42.31	52.54	73.68
Distribución y transporte de gas	33.24	37.66	28.04
Exportadoras e importadoras	43.65 (*)	47.99	52.39 (*)
Financiera y bancos	257.69	64.35	53.62
Frigoríficos	115.63	40.12	43.80 (*)
Gas y petróleo (producción y servicios)	128.53	28.81	46.18
Hotelería y turismo	3.24	3.53 (*)	3.86 (*)
Imprentas y editoriales	450.50	33.94	23.91
Industria pesquera	31.19 (*)	34.29	37.43 (*)
Laboratorios	79.01	55.35	60.43
Lácteos	39.20 (*)	43.10	87.80
Limpieza y cosméticos	158.33	172.63 (*)	188.48 (*)
Máquinas y equipos (bienes de capital)	124.00	48.32	40.18
Metalúrgicas y siderúrgicas	58.72	30.03	31.82
Minero	162.68	22.22	45.05
Molinos y aceites	74.89	24.18	47.77
Neumáticos y productos caucho	28.59	31.17 (*)	28.42
Otros	32.61 (*)	35.85 (*)	39.47
Producción y distribución de energía eléctrica	46.30	18.01	56.93
Productos alimenticios	35.35	42.88	38.85
Productos fotográficos	14.98 (*)	16.47	17.98 (*)
Productos industrias plástica y química	65.97	18.64	77.78
Químicos y petroquímicos	62.35	36.03	36.26
Residuos	6.90 (*)	7.57 (*)	8.32
Servicios de la salud	27.69 (*)	30.43	33.23 (*)
Servicios de seguridad	41.31 (*)	45.41 (*)	50.00
Tabacaleras	23.69 (*)	26.04 (*)	28.67
Tecnología y comunicaciones	61.96 (*)	68.12 (*)	75.00
Telecomunicaciones	58.82	363.10	57.18
Textil y calzado	15.16 (*)	16.67	18.20 (*)
Transporte	60.13	163.61	100.65
Vidrio y mat. construcción	1100.00	22.19	29.56

(*) Estas rentabilidades son estimaciones propias que se han calculado tomando como referencia la rentabilidad del mismo sector correspondiente a otros años afectada por la variación del PBI.

Tabla 10.2 – Rentabilidades del Patrimonio Neto por tipo de actividad expresadas en porcentaje



10.4. Autovalores de las componentes principales

En la Tabla 10.3 se muestran los autovalores de la matriz de covarianzas **C** correspondientes a cada uno de las componentes principales. Nótese que las dos primeras componentes contienen aproximadamente el 93% de la varianza total de la muestra.

<i>Autovalores Varianza Acumulada</i>			<i>Autovalores Varianza Acumulada</i>		
F1	17862.89	84.69%	F23	0.33	99.99%
F2	1792.07	93.19%	F24	0.29	100.00%
F3	576.15	95.92%	F25	0.21	100.00%
F4	455.55	98.08%	F26	0.18	100.00%
F5	134.53	98.72%	F27	0.15	100.00%
F6	88.47	99.14%	F28	0.11	100.00%
F7	62.48	99.44%	F29	0.09	100.00%
F8	50.57	99.68%	F30	0.06	100.00%
F9	19.58	99.77%	F31	0.06	100.00%
F10	11.70	99.82%	F32	0.04	100.00%
F11	9.40	99.87%	F33	0.03	100.00%
F12	5.51	99.89%	F34	0.02	100.00%
F13	4.74	99.92%	F35	0.02	100.00%
F14	4.25	99.94%	F36	0.01	100.00%
F15	3.89	99.96%	F37	0.01	100.00%
F16	2.13	99.97%	F38	0.01	100.00%
F17	1.94	99.97%	F39	0.01	100.00%
F18	1.11	99.98%	F40	0.00	100.00%
F19	0.95	99.98%	F41	0.00	100.00%
F20	0.71	99.99%	F42	0.00	100.00%
F21	0.50	99.99%	F43	0.00	100.00%
F22	0.39	99.99%	F44	0.00	100.00%
			F45	0.00	100.00%

Tabla 10.3 – Autovalores de la matriz de covarianzas

10.5. Valores de las componentes principales retenidas

En la Tabla 10.4 se muestran los valores que toman las componentes principales retenidas en la muestra. A través de ellas, es posible representar a las empresas en un gráfico que tiene como abscisas y ordenadas a dichas componentes.



<i>Empresa N°</i>	<i>F1</i>	<i>F2</i>	<i>Empresa N°</i>	<i>F1</i>	<i>F2</i>
1	2.89	-8.73	44	-21.32	-11.43
2	-7.46	-13.63	45	-24.91	-10.17
3	-29.61	-9.03	46	-29.61	-8.20
4	-28.65	-9.93	47	-27.52	-12.12
5	-25.48	-11.18	48	-27.59	-8.34
6	-28.02	-9.82	49	-19.58	56.95
7	-24.38	-11.21	50	-27.25	-2.63
8	-29.96	-23.40	51	-10.80	348.78
9	-25.94	-9.77	52	-28.00	77.40
10	6.56	-29.62	53	83.18	-0.38
11	-27.35	-12.27	54	22.53	28.75
12	-13.12	-8.80	55	-24.40	-7.53
13	-26.62	-10.99	56	-19.52	-4.45
14	-27.68	-12.63	57	-16.61	-0.07
15	-21.24	-10.02	58	-27.95	-8.51
16	-9.81	-6.73	59	-29.01	-10.73
17	-29.29	-9.85	60	-29.00	-8.98
18	-28.17	-17.67	61	-18.88	-11.57
19	-29.11	-21.38	62	1,168.66	-28.96
20	21.88	-6.99	63	143.16	65.74
21	-27.91	-9.38	64	-9.03	5.18
22	-23.66	-11.65	65	-23.78	-24.24
23	-29.71	-10.96	66	-29.17	-8.94
24	-28.14	-11.62	67	-18.11	-2.91
25	-29.48	-9.49	68	-28.57	-11.15
26	44.80	16.75	69	-28.76	65.83
27	-26.62	-11.21	70	-25.86	-13.43
28	20.48	-6.61	71	-29.66	-7.74
29	-26.89	-11.24	72	-29.15	-9.30
30	-16.64	-8.22	73	-23.67	-8.33
31	-26.94	-10.14	74	-30.66	-3.28
32	-28.95	-10.66	75	-14.80	-5.93
33	49.83	0.62	76	-26.39	6.42
34	-18.95	-21.99	77	-28.18	17.90
35	-28.73	-10.61	78	-26.25	-8.48
36	-28.63	-9.81	79	-29.34	-2.67
37	-29.95	-9.09	80	-27.60	-7.34
38	-27.47	-11.72	81	-29.40	-7.81
39	-25.99	-10.93	82	296.70	55.26
40	-18.58	-7.64	83	-29.29	-10.68
41	-22.00	-14.59	84	0.38	-7.52
42	-28.22	-11.83	85	-29.25	5.02
43	-29.07	-8.22	86	-27.74	4.48

Tabla 10.4 – Valores de las componentes en la muestra



10.6. Correlaciones entre las componentes principales y las variables

En la Tabla 10.5 se muestran las correlaciones entre las componentes principales retenidas y cada una de las variables originales. Se han destacado aquellas variables sobre las cuales las componentes cargan significativamente.

Variable	F1	F2	Variable	F1	F2
X1	-0.0450	0.0003	X23	-0.0032	-0.1069
X2	0.0092	-0.1531	X24	-0.0039	-0.0835
X3	-0.0218	0.0707	X25	0.0158	-0.1549
X4	-0.0395	0.0364	X26	0.0056	-0.1566
X5	-0.0744	-0.0786	X27	0.0149	-0.1681
X6	-0.0150	-0.0299	X28	-0.0453	-0.1014
X7	-0.0251	-0.0285	X29	-0.0167	-0.4457
X8	0.0005	0.0194	X30	-0.0304	-0.3643
X9	0.0031	-0.0038	X31	-0.0262	0.0885
X10	-0.0297	0.1808	X32	-0.0842	-0.0678
X11	-0.0395	0.1606	X33	-0.0033	-0.9688
X12	-0.0299	-0.0295	X34	-0.0819	0.2591
X13	-0.0203	-0.0365	X35	-0.0251	0.1044
X14	0.9984	0.0221	X36	-0.0677	-0.0351
X15	-0.0457	0.1748	X37	-0.0223	0.1054
X16	-0.0357	0.1921	X38	0.0124	0.2825
X17	-0.0323	-0.0277	X39	-0.0323	0.5329
X18	-0.0480	-0.0363	X40	-0.0752	-0.0004
X19	-0.0285	-0.2219	X41	0.8628	-0.2082
X20	-0.0700	-0.3777	X42	0.8120	-0.2345
X21	0.2276	0.0313	X43	-0.0141	-0.7122
X22	-0.0650	-0.0932	X44	-0.0294	-0.4114
			X45	-0.0143	-0.9859

Tabla 10.5 – Correlaciones entre las variables y las componentes principales

10.7. Modelos de regresión logística evaluados

En la Tabla 10.6 se muestran los 57 modelos de regresión logística desarrollados a partir de las variables originales junto con el coeficiente de determinación *Clásico* y de *Nagelkerke*.



Modelo N°	Variables consideradas					$R^2_{Clásico}$	$R^2_{Nagelkerke}$
1	X16	X19	X21	X8	X44	0.57	0.67
2	X16	X29	X21	X23	X44	0.58	0.66
3	X16	X29	X5	X21	X44	0.55	0.65
4	X16	X29	X5	X21	X23	0.56	0.65
5	X16	X29	X26	X23	X44	0.55	0.65
6	X16	X25	X29	X8	X43	0.53	0.64
7	X16	X25	X29	X8	X21	0.54	0.64
8	X16	X29	X5	X23	X44	0.53	0.64
9	X16	X29	X5	X44	X17	0.50	0.64
10	X16	X29	X26	X24	X44	0.53	0.64
11	X16	X29	X21	X24	X44	0.57	0.64
12	X16	X29	X5	X21	X17	0.52	0.63
13	X16	X19	X21	X2	X44	0.54	0.63
14	X16	X25	X8	X29	X19	0.52	0.62
15	X16	X29	X5	X21	X43	0.52	0.62
16	X16	X29	X5	X44	X42	0.50	0.62
17	X16	X29	X21	X44	-	0.54	0.62
18	X16	X19	X21	X3	X44	0.54	0.62
19	X16	X25	X8	X29	-	0.51	0.61
20	X16	X25	X29	X5	X21	0.52	0.61
21	X16	X29	X5	X21	X33	0.51	0.61
22	X16	X29	X5	X21	X32	0.60	0.61
23	X16	X29	X5	X21	X45	0.51	0.61
24	X16	X29	X5	X21	X34	0.51	0.61
25	X16	X29	X5	X23	-	0.50	0.61
26	X16	X29	X5	X44	-	0.48	0.61
27	X16	X29	X5	X44	X20	0.49	0.61
28	X16	X29	X5	X44	X33	0.49	0.61
29	X16	X29	X5	X44	X11	0.51	0.61
30	X16	X29	X2	X21	X43	0.52	0.60
31	X16	X29	X21	X3	X25	0.53	0.60
32	X16	X25	X29	X21	X3	0.53	0.60
33	X16	X25	X29	X21	X2	0.52	0.60
34	X16	X29	X5	X21	-	0.50	0.60
35	X16	X29	X5	X21	X42	0.50	0.60
36	X16	X29	X5	X21	X26	0.51	0.60
37	X16	X29	X21	X23	-	0.54	0.60
38	X16	X3	X25	X33	X8	0.49	0.59
39	X16	X25	X29	X5	-	0.48	0.58
40	X16	X25	X29	X5	X20	0.48	0.58
41	X16	X3	X25	X33	X29	0.49	0.57
42	X15	X29	X2	X43	X21	0.50	0.57
43	X16	X29	X8	-	-	0.47	0.57
44	X16	X25	X29	X3	-	0.49	0.57
45	X16	X25	X29	X2	-	0.49	0.57
46	X16	X29	X5	-	-	0.45	0.57
47	X16	X25	X29	-	-	0.48	0.56
48	X16	X29	X21	-	-	0.49	0.56
49	X16	X3	X25	X33	X19	0.44	0.55
50	X16	X29	X3	-	-	0.47	0.55
51	X16	X25	X3	-	-	0.41	0.53
52	X45	X43	X33	X39	-	0.28	0.36
53	X45	X43	X39	-	-	0.27	0.36
54	X14	X41	X42	X45	X33	0.27	0.34
55	X45	X33	X43	-	-	0.25	0.34
56	X45	X43	-	-	-	0.24	0.33
57	X14	X41	X42	-	-	0.12	0.14

Tabla 10.6 – Modelos de regresión evaluados



10.8. Composición detallada de la nueva muestra de empresas

En la Tabla 10.7 se muestra un detalle de la nueva muestra utilizada para evaluar la efectividad en el reconocimiento de las empresas con problemas financieros cuando se aplica el análisis de componentes principales y el de regresión logística en forma conjunta.

<i>Empresa N°</i>	<i>Grupo N°</i>	<i>Nombre</i>	<i>Periodo Analizado</i>	<i>Actividad desarrollada</i>
1	1	Patricios S.A.	2004-2003	Productos industrias plástica y química
2	1	Fiplasto S.A.	2005-2004	Exportadoras e importadoras
3	1	Grupo Inplast S.A.	2003-2002	Agrícolas y ganaderas
4	1	Grimoldi S.A.	2005-2004	Textil y calzado
5	1	Limpiolux S.A.	2005-2004	Otros
6	1	La Agraria S.A.	2005-2004	Agrícolas y ganaderas
7	1	Amercian Plast S.A.	2004-2003	Productos industrias plástica y química
8	1	Compañía argentina de semillas	2005-2004	Agrícolas y ganaderas
9	1	Schiarre S.A.	2004-2003	Máquinas y equipos (bienes de capital)
10	1	Sports Life S.A.	2005-2004	Comercio minorista
11	1	UOLE S.A.	2005-2004	Artículos del hogar (líneas blanca y marrón)
12	2	Sweet Victorian S.A.	2001-2000	Textil y calzado
13	2	Metcasa Metalúrgica Callegari S.A.	2001-2000	Metalúrgicas y siderúrgicas
14	2	Midan S.A.	2000-1999	Automotores y autopartes

Tabla 10.7 – Empresas relevadas en la nueva muestra de datos

