

**PREDICCIÓN DE LA CRISIS  
BANCARIA EN ESPAÑA  
COMPARACION ENTRE EL  
ANALISIS LOGIT Y EL  
ANALISIS DISCRIMINANTE**

**Joaquina Laffarga Briones  
José Luis Martín Marín  
María José Vazquez Cueto**

Universidad de Sevilla

**Sumario:** 1. Introducción.— 2. Modelos alternativos.— 3. El problema empírico.— 4. Resultados.— 5. Conclusiones.— Bibliografía.

## **1. INTRODUCCION**

Desde los trabajos de Beaver (1966) y Altman (1968) son numerosos los autores que han empleado modelos estadísticos para tratar el problema de predecir la crisis en una empresa. La mayoría de ellos basan sus conclusiones en una técnica de clasificación, prediciendo la pertenencia o no de la empresa en cuestión al grupo de las fallidas en base a su puntuación  $Z$  obtenida a través del modelo lineal discriminante. El propósito de este artículo es el de proponer el modelo logit como alternativa y comparar los dos modelos en un experimento concreto.

## **2. MODELOS ALTERNATIVOS**

Comencemos ofreciendo algunas breves consideraciones relativas a los modelos en estudio.

### **2.1. Análisis discriminante múltiple.**

El análisis lineal discriminante múltiple es una técnica esta-

distica multivariante que nos permite clasificar individuos u objetos en dos o más poblaciones distintas.

Así, si denotamos por  $(X_1, X_2, \dots, X_n)$  al vector de variables independientes o predictores, el método construye la función

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

y asigna al individuo a un grupo u otro dependiendo de que  $Z$  sea mayor o menor que cero.

El principal inconveniente que presenta este modelo es que supone implícitamente que las variables siguen una distribución normal multivariante en cada una de las dos poblaciones, con matrices de varianza-covarianza iguales, diferenciándose solamente en los valores de los vectores medios. Estas hipótesis son difíciles de verificar en un problema práctico. Así, en la mayoría de los problemas financieros, al menos una de las variables (generalmente la variable dependiente) es cualitativa, eliminando la posibilidad de normalidad. En tales casos una técnica que trate específicamente con variables cualitativas, como puede ser el análisis logit, nos parece más oportuna.

## 2.2. Análisis logit.

El análisis logit es una técnica estadística desarrollada para investigar las relaciones entre una variable binaria y una o más variables independientes.

Así, siguiendo con la terminología anterior y codificando los valores de la variable dependiente en 0 y 1, el método nos da para la probabilidad de que  $Y = 1$  la siguiente expresión

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n)}} \quad (2)$$

Obtenidas estas probabilidades, para cada uno de los individuos, el paso siguiente será el de clasificar a los mismos para lo cual el valor  $P(Y = 1) = 0.5$  puede servirnos como punto crítico o punto de corte.

La ventaja que presenta este método es la suavidad de sus hipótesis de partida que son fácilmente verificables en cualquier problema empírico.

## 3. EL PROBLEMA EMPIRICO

Hemos elegido el problema de predicción de la crisis bancaria a través de ratios financieros para comparar las dos técnicas estadísticas a las que se ha hecho mención anteriormente.

La crisis bancaria en España se sitúa cronológicamente en los años 1978 a 1983 cuando entraron en crisis hasta el 47.2% de los bancos españoles existentes en 1978. Este grave problema tiene, paradójicamente, un efecto positivo y es el de proporcionarnos suficiente cantidad de datos contables como para poder abordar un análisis estadístico serio de la cuestión, permitiéndonos establecer una serie de indicadores precoces que puedan servirnos para predecir con suficiente antelación las dificultades financieras en entidades de depósitos. En esta línea se orienta el trabajo presentado por estos mismos autores en la revista ESIC-MARKET (nº 48, Abril-Junio, 1985).

Dado que la base de datos que se ha utilizado en el presente estudio es la misma, a dicho artículo remitimos al lector para una mayor información, limitándonos aquí a una indicación del mismo y a la presentación de las variables que se ha utilizado.

La fuente de datos ha sido el Anuario Estadístico de la Banca Privada publicado por el Consejo Superior Bancario y que recoge los Balances y Cuentas de Pérdidas y Ganancias de los bancos españoles al final de cada ejercicio.

La muestra está constituida por 47 bancos de los cuales 22 son "fracasados" en el sentido dado a continuación.

La variable dependiente o fracaso bancario está considerada como la intervención del banco por parte de las autoridades monetarias y, en concreto, por el Fondo de Garantía de Depósitos. Así, representaremos por  $Y = 1$  el suceso "en el plazo de un año el banco será intervenido por el Fondo de Garantía de Depósitos" y por  $Y = 0$  el suceso contrario.

En cuanto a las variables explicativas, dado que lo que pretendemos es comparar los resultados obtenidos con el análisis discriminante lineal y el análisis logit, más que establecer que ratios financieros pueden dar una mayor información acerca de la crisis, hemos seleccionado cuatro ratios que, a juicio de los investigadores, pueden explicar satisfactoriamente la situación del banco. Estos ratios miden respectivamente la liquidez, estructura del activo, la rentabilidad acumulada y la rentabilidad:

$$R_2 = \frac{\text{Créditos más cartera de valores}}{\text{Activo total}}$$

( $R_2$  es una medida inversa de la liquidez de un banco. Teóricamente, los bancos fracasados deben presentar un valor más alto de  $R_2$ , es decir, deben presentar menos liquidez).

$$R_5 = \frac{\text{Activo fijo}}{\text{Activo total}}$$

( $R_5$  es una medida de las inmovilizaciones en muebles e inmuebles del banco. Los bancos fracasados deben presentar valores más altos de  $R_5$ ).

$$R_9 = \frac{\text{Reservas}}{\text{Pasivo total}}$$

( $R_9$  Representa la rentabilidad acumulada del banco a lo largo del tiempo, después del pago de impuestos y dividendos. Los bancos fracasados deben presentar un  $R_9$  menor).

$$R_{11} = \frac{\text{Beneficio neto antes de impuestos}}{\text{Activo total}}$$

( $R_{11}$  mide la rentabilidad, antes de impuestos, de todo el activo del banco.  $R_{11}$  debe ser menor para los bancos fracasados).

Dividida la muestra de bancos considerada en dos submuestras, correspondientes a los bancos fracasados y saneados respectivamente, se analizan los valores que toma cada ratio en cada uno de ellos y se contrasta su significación mediante un test de igualdad de medias. Para ello se construye el estadístico F que se distribuye según una F de Snedecor con 1 y  $N_1 + N_2 - 2$  grados de libertad, siendo  $N_i$  el nº de observaciones del grupo  $i$ ,  $i = 1, 2$ . Cuanto mayor sea el valor de F mayor es la variación entre grupos para el ratio en cuestión y con mayor propiedad, por tanto, podremos decir que dicho ratio distingue a los grupos.

Los resultados se ofrecen en el cuadro nº 1 donde se indica también la media y desviación típica de cada ratio en cada submuestra.

Cuadro nº 1: Ratios seleccionados:

|          | Bancos saneados |              | Bancos fracasados |              | Test F | Nivel de significación |
|----------|-----------------|--------------|-------------------|--------------|--------|------------------------|
|          | media           | desv. típica | media             | desv. típica |        |                        |
| $R_2$    | 0'346           | 0'100        | 0'414             | 0'124        | 4'24   | <0'05                  |
| $R_5$    | 0'036           | 0'020        | 0'053             | 0'023        | 7'25   | <0'01                  |
| $R_9$    | 0'042           | 0'026        | 0'024             | 0'015        | 7'80   | <0'01                  |
| $R_{11}$ | 0'011           | 0'009        | 0'003             | 0'013        | 18'35  | <0'0001                |

#### 4. RESULTADOS

Una vez definidas las variables que intervienen en el estudio pasemos a relacionarlas por medio del análisis discriminante y el análisis logit.

Los resultados del análisis lineal discriminante se ofrecen en el cuadro nº 2 donde se indican los valores estimados de los coeficientes y sus errores asintóticos estimados. La estimación ha sido realizada utilizando el paquete de programas BMPD, del Departamento de Biomatemáticas de la Universidad de California, que se encuentra disponible en el Centro de Cálculo de la Universidad de Sevilla.

El método utilizado para el estudio de la significación ha sido el de borrado condicional.

Cuadro nº 2: Resultados estimación análisis discriminante.

|                       | Coefficiente | Error asintotico | t-Student |
|-----------------------|--------------|------------------|-----------|
| cte                   | -1'91        |                  |           |
| R <sub>2</sub>        | 4'21**       | 2'008            | 2'132     |
| R <sub>5</sub>        | 26'18***     | 10'188           | 2'610     |
| R <sub>9</sub>        | -16'6        | 11'04            | -1'529    |
| R <sub>11</sub>       | -54'7***     | 20'77            | -2'678    |
| Nº observaciones = 47 |              | Niveles de       | * 10%     |
| Test F = 8'845        |              | significación    | ** 5%     |
|                       |              |                  | *** 1%    |

El cuadro nº 3, análogo al cuadro nº 2 ya explicado, ofrece los resultados de la estimación logit. Las estimaciones han sido obtenidas utilizando el programa "Loglin" desarrollado por Nerlove y Press de la Universidad de Chicago que gentilmente nos lo han proporcionado.

Cuadro nº 3: Resultado estimación logit.

|                             | Coefficiente | Error asintotico | t-Student |
|-----------------------------|--------------|------------------|-----------|
| cte                         | -0'812       |                  |           |
| R <sub>2</sub>              | 4'61         | 3'036            | 1'519     |
| R <sub>5</sub>              | 22'25*       | 13'43            | 1'656     |
| R <sub>9</sub>              | -26'6*       | 14'65            | -1'818    |
| R <sub>11</sub>             | -218'65***   | 87'707           | -2'492    |
| Nº observaciones = 47       |              | Niveles de       | * 10%     |
| Test x <sup>2</sup> = 24-26 |              | significación    | ** 5%     |
|                             |              |                  | *** 1%    |

Una simple revisión de estos cuadros nos muestra como la significación de las variables independientes varía sustancialmente de un modelo a otro. Así el ratio R<sub>2</sub>, ratio de liquidez, resulta significativo, con un nivel de confianza del 96% aproximadamente, en el modelo discriminante y no lo es en el análisis logit. De la misma forma el ratio R<sub>5</sub>, de estructura del activo, tiene un nivel de confianza del 98'8% en el modelo discriminante y solo del 91% en el modelo logit. En cuanto a los ratios R<sub>9</sub> y R<sub>11</sub>, que miden la rentabilidad, los resultados son muy dispares, siendo los niveles de confianza de R<sub>9</sub>, en el modelo discriminante y en el modelo logit, del 87% y del 94% respectivamente. Para R<sub>11</sub> el valor es del 99% en ambos modelos.

En cuanto a la significación de los modelos, que puede ser medida a través del test F en el modelo discriminante y por el test X<sup>2</sup> en el análisis logit, vemos como ambos son suficientemente significativos resultando que el modelo discriminante explica el 99'2% de la variación de la variable dependiente y el 99'5% el modelo logit.

Por otra parte, y para poder determinar que modelo clasifica de forma más correcta a las observaciones, hemos obtenido la puntuación Z del análisis discriminante y la P (Y = 1) del análisis logit para cada banco de la muestra y asignado un banco al grupo de los fracasados si Z > 0 ó P (Y = 1) > 0'5 respectivamente. Los resultados de la clasificación se ofrecen en los cuadros nº 4 y nº 5.

Cuadro nº 4: Porcentaje clasificación correcta. Modelo discriminante.

|          |       | Pronóstico             |     |       |
|----------|-------|------------------------|-----|-------|
|          |       | FRAC                   | SAN | %     |
| Realidad | FRAC  | 18                     | 4   | 81'81 |
|          | SAN   | 2                      | 23  | 92'00 |
|          | TOTAL | 20                     | 27  | 86'91 |
|          |       | FRAC = Banco fracasado |     |       |
|          |       | SAN = Banco saneado    |     |       |

Cuadro nº 5: Porcentaje clasificación correcta. Modelo logit.

|          |       | Pronóstico |     |       |
|----------|-------|------------|-----|-------|
|          |       | FRAC       | SAN | %     |
| Realidad | FRAC  | 20         | 2   | 90'50 |
|          | SAN   | 1          | 24  | 96'00 |
|          | TOTAL | 21         | 26  | 93'40 |

FRAC = Banco fracasado  
SAN = Banco saneado

Estos cuadros muestran claramente como la estimación de los parámetros del modelo logit produce resultados que dominan completamente a los producidos por el modelo discriminante. Se puede observar que, independientemente de que el porcentaje de clasificación correcta total es mucho más alto en el modelo logit, la diferencia en el porcentaje de aciertos de los bancos fracasados es mucho mayor. Lo anterior es un resultado altamente significativo pues parece lógico que el coste de clasificación errónea sea en los bancos fracasados mucho más elevado que en bancos saneados.— (Coste de errores tipo I mayor que coste de errores tipo II).—

## 5. CONCLUSIONES

Como apartado final de este trabajo y a modo de conclusión podemos hacer notar lo siguiente:

1º Siendo la hipótesis de normalidad exigida para la aplicación del modelo discriminante difícilmente verificable en los problemas empíricos de predicción de crisis y dado que en ellos la variable dependiente es cualitativa, parece, a priori, más apropiado la utilización del análisis logit, desarrollado específicamente para tratar con este tipo de variables.

2º Comparando los dos modelos en un problema concreto los resultados vienen a confirmar nuestras suposiciones pues, prescindiendo de cualquier otra consideración y dado que ambos modelos poseen una alta significación estadística, los porcentajes de clasificación correcta en el modelo logit resultan más elevados.

## BIBLIOGRAFIA

- ALTMAN, EDWAR I. (Septiembre 1968): "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy" —The Journal of Finance, Vol. 23, Nº 4.
- BEAVER, WILLIAM H. (1966): "Financial Ratios as Predictors of Failure" Empirical Research in Accounting: Selected Studies. Suplemento al Vol. 5 del Journal of Accounting Research.
- CONSEJO SUPERIOR BANCARIO. (1977 a 1982): Anuario estadístico de la Banca Privada.
- LECHENBRUCH, PETER A. (1975): "Discriminant Analysis". Hafner Press, Nueva York.
- LAFFARGA, JOAQUINA; MARTIN, JOSE LUIS y VAZQUEZ, MARIA JOSE (Junio, 1985): "El Análisis de la Solvencia en las Instituciones Bancarias: Propuesta de una Metodología y Aplicaciones a la Banca española". Esic-Market nº 48. Abril - Junio.
- SINKEY, JOSEPH F. (Marzo, 1975): "A Multivariate Statistical Analysis of the Characteristics of Problem Banks". The Journal of Finance, vol. 30, nº 1.
- TABACHNICK, BARBARA G. y FIDELL, LINDA S. (1983): Using Multivariate Statistics. Harper & Row, Nueva York.