



## Artículos

# Éxito-Fracaso en Empresas Regionales: Aplicación del método "stepwise"

Alonso, Aldo H. \*\*; Legato, Ana María\*; Rébora Alfredo\*

*El objetivo fundamental del presente trabajo es desarrollar y analizar la aplicabilidad para una región (Tandil-Argentina), de un modelo multivariado para clasificar y predecir problemas financieros en PyME, empleando el procedimiento "stepwise" en la metodología del Análisis Discriminante. Para ello se ha modelizado empíricamente empleando 17 índices financieros sobre una muestra de 53 empresas de la región, correspondientes a distintos sectores de actividad y clasificadas en 2 grupos: Empresas Exitosas y Empresas Fallidas. La función discriminante resultante define como mejores variables a efectos de clasificar y discriminar entre ambos tipos de empresas a: Endeudamiento respecto al Activo Total, Inmovilización de Activos y Rotación sobre el Activo Total. Se comprobó que el modelo clasificó correctamente al 81,13% del total de la muestra, porcentaje que evidencia la confiabilidad de las conclusiones que derivan de la aplicación del mismo. Es decir, proporciona una predicción bastante exacta en cuanto a la clasificación de las empresas, o sea, una estimación optimista de la calidad de la discriminación.*

### INTRODUCCION

El análisis de ratios que distingue, por ejemplo, entre ratios de liquidez, endeudamiento, rentabilidad, actividad, etc., se ha efectuado al margen de un contexto definido de la teoría de toma de decisiones. La crítica más importante deriva del hecho que el mismo ha ignorado desarrollos registrados en campos como el de la economía y las finanzas. En el análisis financiero moderno se pregona en que las técnicas de análisis financiero, como el caso de los ratios, deben insertarse en un marco de modelos decisorios. Por tanto, una diferencia significativa entre el enfoque tradicional y el moderno es que, en el primero, el análisis se efectúa al margen de un ámbito bien definido de teoría decisoria, algo que el segundo entiende esencial.

En esta línea de pensamiento se vienen desarrollando esfuerzos por introducir el análisis de ratios (entre otras aplicaciones) dentro del modelo de predicción de las ganancias de las empresas, el crecimiento de éstas y la constitución de carteras eficientes.

Un área donde se ha desarrollado el enfoque moderno, es el de la predicción de problemas financieros en las empresas. Al respecto, dichos ratios han sido insertados en técnicas estadísticas multivariadas, como es el caso del análisis discriminante.

El tema de la predicción de problemas financieros a través del análisis discriminante, ha merecido ser objeto de preponderante atención por parte de distinguidos especialistas. Entre ellos cabe mencionar a

Beaver (1967) quien comenzó sus estudios considerando separadamente varios ratios financieros y el valor individual a partir del cual puede considerarse una situación crítica; Altman (1968) en el desarrollo del modelo de puntación Z y posteriormente (1977) con el modelo de riesgo crediticio Zeta<sub>R</sub>. Ambos modelos están siendo utilizados por los profesionales en el área en todo el mundo. Siguió el camino de Altman, Spingate (1978) en Canadá, empleó el método "paso a paso" a un conjunto de 40 empresas logrando una proporción de exactitud del 92,5%. En la universidad de Quebec, Legault (1987) analizó 30 proporciones financieras en una muestra de 173 negocios industriales de Canadá y obtuvo el modelo conocido como Ca\_Cuenta con una confiabilidad del 83%. O'Neill y otros (2002) aplicaron la metodología a un grupo de 178 empresas de US con una confiabilidad del 80%.

Como ejemplo de modelos para países de menor desarrollo pueden mencionarse, entre otros a Pascale (1984), quien trabajó con datos de empresas de la industria manufacturera en Uruguay logrando 91 % de clasificación correcta. Mariaca (2002) se basó en una muestra perteneciente al sector bancario de Bolivia, el porcentaje de clasificación correcta fue del 69% y Delfino (2002) en un trabajo sobre la predicción del default de entidades financieras en Argentina.

### ANÁLISIS DISCRIMINANTE: EL MODELO

#### MARCO TEORICO

El análisis discriminante es una técnica de metodología estadística usada para clasificar una observación en uno o varios agrupamientos a priori que dependen de las características individuales de la

(\*) Universidad Nacional del Centro de la Pcia de Bs.As. Facultad de Ciencias Económicas, Ctr de Estudios en Administración y Economía, Pto N° 399, C7000 Tandil-Argentina.

(\*\*) Universidad Nacional de La Plata, Facultad de Ciencias Económicas, Calle 6 N° 777, (1900) La Plata-Argentina.

observación. Después de haberse determinado los grupos se reúne la información para los elementos en dichos grupos. La técnica en cuestión, en su forma más simple, trata de obtener una combinación lineal de esas características, tal que sea la que "mejor" discrimina entre los grupos. En el gráfico siguiente puede apreciarse, en general el efecto de la función *discriminante*, cuando en la población fueron definidos dos grupos de empresas.

•Gráfico 1 : Análisis Discriminante en dos Grupos.



La función discriminante, objetivo primordial de este trabajo, en general adopta la forma

$$Z = a_0 + a_1X_1 + a_2X_2 + \dots + a_kX_k \quad (1)$$

Donde: Z = valor discriminante calculado para cada empresa particular;  
 $a_i$  = coeficientes discriminantes, con  $i = 1, 2, \dots, k$ ;  $x_i$  = variables independientes (índices de balances), con  $i = 1, 2, \dots, k$ ; transforma los valores de las variables individuales en una sola puntuación discriminante,  $\varphi$  valor Z, que luego es usado para clasificar el objeto.

#### Formulación matemática del modelo

El análisis discriminante se puede abordar desde diversos aspectos, pero en todos los casos, para que la función discriminante lineal sea óptima, es decir, minimice la probabilidad de clasificación errónea, se deben formular hipótesis sobre los datos.

Ellas son:

- a. Cada grupo debe ser una muestra de una población normal multidimensional
- b. Las matrices de varianzas-covarianzas de los dos grupos deben ser iguales; sea  $k$  el número de grupos;  $P$  el conjunto de variables observadas y  $p$  el número de variables discriminantes;  $n_i$  número de individuos por grupo; entonces  $k \geq 2, 0 < p \leq n_i - 2$
- c. Ninguna variable discriminante es combinación lineal de otras variables discriminantes.

Teniendo en cuenta estos supuestos básicos, en el presente trabajo se utilizará el procedimiento "paso a paso" o "stepwise" para la selección de las variables a incluir en el modelo.

#### GENERALIDADES DEL METODO "stepwise"

El procedimiento "paso a paso" consiste en eliminar cierto número de variables, para considerar solamente aquellas que probaron ser las más significativas desde el punto de vista de la discriminación.

Más precisamente, siendo  $p$  variables medidas en un conjunto  $E$  de  $n$  individuos (empresas), se buscará primero la variable más discriminante en función de un criterio determinado. Después la variable que entre las  $(p - 1)$  que quedan, junto con la variable que ya fue seleccionada, genera mejor discriminación empleando ese mismo criterio, luego la variable que entre las  $(p - 2)$  que quedan, junto con las dos variables ya seleccionadas, produce la mejor discriminación y así sucesivamente hasta que la selección de la última variable que contribuya a la discriminación en función del criterio empleado.

Este procedimiento requiere que se elija un criterio de la selección para determinar variables sucesivas. A tal efecto en base a la metodología desarrollada por Langrand (1997) se optó por el Lambda de Wilks.

Lambda de Wilks es una medida multivariante de las diferencias entre grupos sobre varias variables discriminantes. Valores de Lambda próximos a cero indican alta discriminación, es decir, los grupos centroides están separados y son muy distintos en relación a la dispersión que existe dentro de los grupos.

A medida que el Lambda va creciendo, indica en forma progresiva, menos discriminación. Si Lambda vale uno, indica que los grupos centroides son idénticos, no existen diferencias.

Simultáneamente al lambda, el método "stepwise" se apoya en dos estadísticos  $F$  "to enter" y  $F$  "to remove" (que pueden traducirse como  $F$  "de entrada" y  $F$  "de salida"), mediante los cuales una variable cualquiera va a ser incluida o excluida de la función discriminante.

Suponiendo que hay  $q < p$  variables en el análisis, el estadístico  $F$  "de entrada" de  $X_i$ , que se denota  $F_i$ , permite realizar un contraste acerca de la igualdad de las medias de la variable  $X_i$  en los  $k$  grupos, considerando las variables que están ya en el análisis:

Un valor grande de  $F_i$  indica que la variable  $X_i$  junto con las  $q$  que ya había contribuye "algo" a la separación de los grupos. Es decir, al incluir  $X_i$  la separación de los grupos aumentará significativamente.

Por otro lado, suponiendo nuevamente que hay  $q < m$  variables en el análisis, mediante el estadístico "F de salida" de  $X_i$ , que se denota  $F_i$ , se contrasta entonces la hipótesis de igualdad de  $k$  medias  $X_i$  cuando las otras  $q-1$  variables se parcializan.  $F_i$  es el grado en que  $X_i$  junto con las  $q-1$  variables contribuye a la separación de los grupos.

Un valor de  $F_i$  "grande" indica  $X_i$  junto con las restantes  $q-1$  variables contribuye "algo" a la separación de los grupos decrecerá significativamente. Cada valor  $F_i$ , siendo  $F$  el estadístico de Fisher-Snedecor, lleva asociado un valor del estadístico Lambda de Wilks.

## METODOLOGIA

### Selección de la muestra

La selección de la muestra se realizó de acuerdo con los objetivos del trabajo, es decir, considerar dos grupos de empresas, designados como empresas fallidas o con problemas y empresas exitosas o sin problemas.

La definición de empresa fallida o con serios problemas financieros se efectuó sobre la base de la comprobación de acontecimientos que pusieran de relieve situaciones críticas, como convocatorias, quiebras, arreglos con bancos u otros entes financiadores, etc. El número de empresas seleccionadas dentro del grupo fue finalmente de veintiseis (26). Una vez seleccionada la muestra de empresas con problemas, se procedió a elegir la correspondiente empresa sin problemas. Así, por cada empresa fallida, se trató de encontrar una empresa sana perteneciente al mismo sector económico y cuyos datos perteneciesen al mismo período considerado para la empresa fallida. El número de empresas seleccionadas en este caso fue de veintisiete (27).

En consideración a la metodología del trabajo (trabajar con proporciones o ratios), se recogieron como variables para ingresar en el análisis, a diez y siete de ellos, que en parte surgen del aporte de la literatura especializada y, también de la experiencia recogida en la realidad de algunos países menos desarrollados que sufren procesos inflacionarios. Los ratios se clasificaron en 4 grupos: Ratios de Rentabilidad; de Liquidez; de Administración de Activos y de Endeudamiento.

### Desarrollo del método y obtención del modelo

Teniendo en cuenta la filosofía del procedimiento paso a paso, las variables seleccionadas para "armar" el modelo surgen de la siguiente sucesión de pasos:

#### Primer Paso:

Para la muestra de empresas, la variable que representa el menor lambda de Wilks, es Endeudamiento sobre Activo Total, con un Lambda = 0,657, es por lo tanto la candidata a ser seleccionada en el primer paso. (Tabla 1). En esta tabla y en las 3 siguientes no se presenta la información completa por falta de espacio. La información de la totalidad de las tablas corresponde a resultados del comando SPSS, Análisis Discriminante para dos grupos.

Pero, que una variable sea candidata a la selección no implica que resulte seleccionada, habrá que establecer un criterio para determinar si la información aportada por dicha variable es significativa; de no serlo, se la eliminará. A tal efecto se deben considerar los valores de F de entrada, F de salida y Tolerancia. Debido a las características de los datos y, después de varios ensayos, los valores de F de entrada y F de salida quedaron establecidos en 2,25 y 1,27 respectivamente (los valores estándares son 3,84 y 2,71). El F "de entrada" de la razón end. s act. total supera el límite.

#### Segundo paso:

En el segundo paso, de entre las restantes variables independientes, la

candidata a ser seleccionada es la que presenta el mínimo valor del Lambda, según el esquema de la Tabla 2.

Es así como la segunda variable a seleccionar por su poder discriminante resultó ser Inmovilización sobre Activos con una F de entrada igual a 2,554 (mayor que el 2,25 establecido). En consecuencia no solo es la candidata sino que será seleccionada en el segundo paso y no se podrá eliminar dado que su F de salida (Tabla 5) es 2,554 mayor que el 1,27 establecido.

#### Tercer paso:

Una vez comprobado que ninguna variable puede ser eliminada, el siguiente paso será comprobar si la F de entrada correspondiente a la variable candidata de ser seleccionada, la que proporcione el mayor valor Lambda de Wilks, es mayor que 2,25. Al igual que en los pasos anteriores, se analizan los estadísticos F de entrada y F de salida correspondientes, según Tabla 3, la candidata es Rot. Act. Tot, pues es la que presenta el menor Lambda (0,628) y el F de entrada es superior a 2,25, luego será la seleccionada en el paso actual.

#### Cuarto Paso:

Con el mismo razonamiento que en los tres pasos previos, analizamos la información contenida en la Tabla 4.

De entre las variables a ser seleccionadas en el tercer paso, la candidata es Ganancia operativa, pero como su F de entrada igual a 1,098 es menor que el límite establecido, no cumple con las condiciones fijadas, por lo tanto no puede ser incluida en el modelo.

Tabla 1. Variables seleccionadas para el análisis

Paso 1	Tolerancia	Tolerancia mínima	F de entrada	Lambda de Wilks
End. s Act. Total	1,000	1,000	26,627	0,657

Tabla 2. Variables seleccionadas para el análisis

Paso 2	Tolerancia	Tolerancia mínima	F de entrada	Lambda de Wilks
Inm. Act	0,992	0,992	2,554	0,625

Tabla 3. Variables seleccionadas para el análisis

Paso 3	Tolerancia	Tolerancia mínima	F de entrada	Lambda de Wilks
Rot. Act. Total	0,962	0,952	3,255	0,586

Tabla 4. Variables seleccionadas para el análisis

Paso 4	Tolerancia	Tolerancia mínima	F de entrada	Lambda de Wilks
Gan Op	0,991	0,949	1,098	0,573

Tabla 5. Variables incluidas en el análisis

Paso	Tolerancia	F que eliminar	Lambda de Wilks
1 End. s/Act. Total	1,000	26,627	
2 End. s/Act. Total	0,992	26,526	0,957
Inm. Act.	0,992	2,554	0,657
3 End. s/Act. Total	0,989	25,294	0,889
Inm. Act.	0,956	3,470	0,628
Rot. Act. Total	0,962	3,255	0,625

Tabla 6. Bondad del modelo para la clasificación

Empresas pertenecientes actualmente al grupo	Empresas clasificadas dentro del grupo		Porcentaje correcto
	Con problemas	Sin problemas	
Con problemas (CPF)	22	4	84,61
Sin problemas (SPF)	6	21	77,77
Total	28	25	81,13

En consecuencia, dado que ninguna variable más puede ser eliminada o seleccionada, el proceso finaliza

**RESUMEN DE VARIABLES Y PASOS**

En la Tabla 5 puede observarse la fuerte disminución que ha experimentado el Lambda de Wilks desde el primer paso 0,957 hasta el tercero en el que su valor es 0,625. Por lo tanto, al representar al conjunto de las empresas en el espacio generado por las tres variables End.s/Act.Tot., Inmv.Act. y Rot.Act.Tot., los grupos estarán bien separados, y por ende, es posible que las funciones discriminantes que se extraigan a partir de dicho espacio, proporcionen un alto porcentaje de clasificaciones correctas en los grupos. Clasificación comprobada mediante el indicador tolerancia, que se define como

$$Tol_j = 1 - R_j^2 \tag{2}$$

donde  $R^2$  es el coeficiente de correlación múltiple entre  $X_j$  y las variables  $X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_p$ .

Si el valor de la tolerancia es igual a 0, la variable  $X_j$  será una combinación lineal de las restantes variables y por lo tanto excluida de la ecuación. El valor crítico establecido para la tolerancia es igual a 0,001.

**RESULTADOS Y DISCUSION**

**La función discriminante**

Con los datos suministrados y el proceso iterativo señalado, la función discriminante que finalmente se obtuvo es:

$$Z = 0,912P/AT + 0,409 \text{ Act.Inm.} + 0,396 \text{ Vtas./AT} \tag{3}$$

donde:  $P/AT$  = Endeudamiento respecto del Activo Total;  $\text{Act. Inm}$  = Inmovilización de Activos;  $\text{Vtas./AT}$  = Relación Ventas sobre Activo Total.

**Validación de los resultados de la clasificación:**

La clasificación de las empresas se realiza a través de la función discriminante (3) y la técnica utilizada se basa en la regla de Bayes. La probabilidad estimada de que una empresa  $i$ , con score discriminante  $Z_i$  pertenezca al grupo  $j = 1, 2$  se denota como  $P(G_j | D)$  y se obtiene con la expresión:

$$P(G_j | D) = \frac{P(D | G_j) P(G_j)}{\sum_{j=1}^2 P(D | G_j) P(G_j)} \tag{4}$$

Donde:

$D$  scores proporcionados por  $Z$ ,  $P(G_j)$  es la probabilidad a priori de pertenecer al grupo  $j$  y  $P(D|G_j)$  es la probabilidad de que, supuesto que la empresa pertenece al grupo, sus scores en la función discriminante pertenezcan efectivamente a ella.

Una empresa será clasificada en el grupo para el cual la probabilidad a posteriori sea máxima, es decir, será clasificada en  $G_j$  si:

$$P(G_j | D) = \max P(G_1 | D), P(G_2 | D) \tag{5}$$

A partir de probabilidades a priori y de las puntuaciones discriminantes, la clasificación resumen de resultados se exponen en la Tabla 6.

De acuerdo con la información resumida, se desprende que el modelo proporciona el 81,13% de clasificación correcta. Del total de empresas  $E$  originales solamente un 22% fue considerada por el modelo como definida erróneamente, en tanto que del grupo de las empresas  $F$  el 84,61% fue clasificada correctamente, de acuerdo a los indicadores seleccionados considerados de mayor poder discriminante.

**PREDICCIÓN**

Hasta el momento se ha comprobado que, mediante la función  $Z$  construida a partir de los valores de los ratios  $P/AT$ ,  $\text{Act.Inm.}$  y  $\text{Vtas/AT}$ , el porcentaje correcto de las clasificaciones es elevado. Por lo tanto es de esperar que dicha función proporcione buenos resultados a la hora de predecir.

Análisis posteriores, y en esto se centra la innovación específica del trabajo, respecto a otras investigaciones anteriores en las cuales se trabajó con una sola función  $Z$  demostraron la ventaja en cuanto a la asignación de una empresa en particular a uno de los dos grupos, de trabajar con dos funciones discriminantes surgidas del mismo análisis, (previo a la etapa de estandarización) una correspondiente al grupo de las exitosas ( $Z_E$ ) y otra al grupo de las fallidas ( $Z_F$ ).

De este modo, las funciones discriminantes adoptan la forma:

$$ZF = -4,292 + 8,721P/AT + 8,511Act.Inm + 1,076 Vtas./AT \quad (6)$$

$$ZF = -9,713 + 15,709P/AT + 11,406Act.Inm. + 1,648 Vtas./AT \quad (7)$$

Si se compara la información contenida en (6) y (7) con la expresada en la ecuación (3), puede observarse que las variables discriminantes son las mismas pero con pesos diferentes y término independiente no nulo debido a que dichas funciones se obtuvieron partiendo de los ratios originales y no de los estandarizados.

En este caso, se calculan los valores Z para una empresa en particular, dicha empresa se incorpora al grupo cuya función discriminante tenga el valor máximo. Y es esta la forma de predicción no especificada en la bibliografía correspondiente, un modo más directo de clasificar a las nuevas empresas dado que no se "evalúa" la zona de indiferencia.

## CONCLUSIONES

El desarrollo presentado prueba el potencial analítico del análisis empleado como asimismo el valor práctico inherente al uso de criterios financieros.

Los resultados obtenidos por la función discriminante pueden considerarse suficientemente buenos, dado que el porcentaje de bien clasificados es del 81,13 %, es decir, una proporción similar a la conseguida por otros modelos referidos en el presente trabajo. Lo habilita, en consecuencia, como instrumento capaz de aportes de importancia en la instancia de tener que realizar un diagnóstico financiero. Por consiguiente, la utilidad de la información contable para identificar situaciones de insolvencia y, por lo tanto, para la toma de decisiones relacionadas con ésta ha sido verificada.

Tal entendimiento, no debe pretenderse que reemplace a los otros instrumentos del análisis y la praxis en materia financiera.

En resumen, este indicador se suma a los elementos de diagnóstico ya existentes, enriqueciendo el análisis teórico de los problemas y el instrumental a través de cuyo uso se busca optimizar los objetivos empresarios, a partir de su supervivencia.

## BIBLIOGRAFIA

- \*Alonso A. y Legato A., "Factores asociados con el éxito-fracaso de empresas y microemprendimientos - Análisis de variables cuali y cuantitativas". *Disertación XIII Jornadas Nacionales de Administración Financiera, Argentina (1998)*.
- \*Alberici, A., "Analisi dei bilanci e previsione delle insolvenze", *ISEDI: Istituto Editoriale Internazionale, Milano (1975)*.
- \*Altman, E., "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance, USA (1968)*.
- \*Altman, E., Haldeman, R. y Narayanan, P., "ZETA Analysis: a new model to identify bankruptcy risk of corporations", *The Journal of Banking and Finance, USA (1977)*.
- \*Audresch, D. B. and Vivarelli, M., "New firm start-ups in Italy", *Discussion paper series, # 864, Centre for Economic Policy Research, November (1993)*.
- \*Beaver, W.H., "Financial ratios as predictors of failure", *Journal of Accounting Research, 83-86 (1967)*
- \*Damodaran, A., "Corporate finance: Theory and practice". Ed. J. Wiley & Sons, USA (1997).
- \*Delfino A., "Predicción de default de entidades financieras", *Banco Central Republica Argentina (2002)*.
- \*Langrand, C., "Análisis Discriminante", *Université des Sciences et Technologies de Lille, France, (1997)*
- \*Luerti A., "La previsione dello stato di insolvenza delle aziende (Credit Scoring)", *Pirola Editore, Milano (1989)*.
- \*Lebart L., Piron M. y Morineau A., "Statistique exploratoire multidimensionnelle", *Dunod, Paris, (1995)*.
- \*Legault J., "CA-Score, a warning system for small business failures" *Bilanas : 29-31 (1987)*.
- \*López, A., "Un Modelo de la predicción de la insolvencia empresarial en el sector textil de Barcelona". *Documents N° 2001/7, Universitat Autònoma de Barcelona, España (2001)*
- \*Mariaca R., "Predicción de problemas de crisis y continuidad en empresas bancarias". *Documento de trabajo N° 11/12, Instituto de Investigaciones Socio Económicas, Bolivia (2002)*.
- \*O'Neill M., Brabazon A., Ryan C. Y Matthews R., "Gramatical evolution and corporate failure prediction", *IEEE Trans. Evolutionary Computation (2002)*.
- \*Pascale R., "Un Modelo multivariante para predecir problemas financieros en empresas. El caso Uruguayo" *Administración de Empresas: 177 (1), 713-723 (1984)*.
- \*Springate G., "Predicting the Possibility of Failure in a Canadian Firm", *Unpublished MBA Research Project Simon Fraser University (1978)*.
- \*Van Nezel M., Sere K., Laintinen T., "Choosin Ban Kruptcy Predictors Using Discriminant Análisis. Logist Analysis and Genetic Algorithms", *Department of Accounting and Finance, FIN-20520 Turku, Finlnd (2000)*.