



Instituto Superior de Economia e Gestão

## SINAIS DE ALERTA DE CRISE EMPRESARIAL: MODELO DE REGRESSÃO LOGIT BASEADO EM DADOS PORTUGUESES

Lúcia Lima Rodrigues

Escola de Economia e Gestão, Universidade do Minho

Ana Maria Bandeira

Instituto Superior de Contabilidade e Administração do Porto

### Resumo

Nas últimas três décadas têm vindo a ser desenvolvidos modelos de previsão de falência multidimensionais, ou seja, modelos que analisam simultaneamente um número de características chave, tais como rentabilidade, solvabilidade e liquidez. O primeiro procedimento a ser usado foi a análise discriminante e, mais recentemente, os modelos de regressão *logit* e *probit* têm vindo a vulgarizar-se dado terem a vantagem de serem não lineares.

Após uma revisão selectiva da literatura, neste trabalho apresentamos um modelo de regressão *logit* com o objectivo de obter sinais de alerta de crise empresarial no nosso país.

Palavras-chave: Modelos de previsão de falências; Modelos multidimensionais; análise discriminante; modelo de regressão *logit*; caso Português.

### 1. INTRODUÇÃO

A falência empresarial constitui uma preocupação para os gestores, accionistas (sócios), credores, trabalhadores, revisores oficiais de contas e outros. Por isso, detectar rapidamente na empresa sinais de crise empresarial tornou-se muito importante para a Contabilidade e Finanças. Desde o estudo embrionário de BEAVER (1966), que temos vindo a assistir ao aparecimento de outros trabalhos que tentam desenvolver modelos apropriados de previsão de falências.

A construção de sinais de alerta de crise empresarial pode ser efectuada numa abordagem unidimensional, baseada em indicadores isolados, seguindo a linha de orientação do trabalho de BEAVER (1966). Contudo, dadas as limitações deste tipo de análise, a revisão selectiva da literatura, efectuada na secção 2., começa com a análise multidimensional baseada na combinação de diversos indicadores, iniciada por ALTMAN (1968) até chegarmos aos modelos de regressão *logit*.

Procurando contribuir com algo de novo, na secção 3. estimamos uma função estocástica (modelo *logit*) que, para características conhecidas de uma empresa portuguesa, nos fornece a respectiva probabilidade de sobrevivência. Ou seja, os resultados que se obtêm deverão permitir uma actuação prévia ao desencadear de fenómenos de crise, de modo a procurar evitá-los.

## 2. REVISÃO SELECTIVA DA LITERATURA

### 2.1. Análise discriminante múltipla

#### 2.1.1. O contributo de ALTMAN (1968)

O trabalho pioneiro de ALTMAN (1968) abriu caminho à aplicação de métodos estatísticos que permitem obter um conjunto de resultados e instrumentos operacionais que possibilitam a identificação das empresas para as quais a probabilidade de ocorrência de dificuldades é mais elevada.

ALTMAN (1968) utilizou pela primeira vez a análise discriminante múltipla com o objectivo de confrontar dois grupos de empresas comparáveis, quer por sector de actividade, quer por capacidade produtiva, definidos previamente, tendo em conta um perfil de comportamento, expresso através de vários rácios.

Basicamente, de acordo com este método estatístico começa-se por escolher, de um conjunto de rácios previamente seleccionado, o rácio que mais evidencia as divergências de comportamento entre os grupos de empresas. Em seguida, entre os restantes rácios, escolhe-se aquele que, conjuntamente com o rácio inicialmente seleccionado, melhor evidencia as divergências entre os comportamentos dos grupos das empresas seleccionadas. Seguindo esta metodologia, formula-se a combinação linear dos rácios previamente seleccionados, que torna mais significativa a diferença entre os comportamentos dos grupos de empresas.

Mais especificamente, podemos dizer que ALTMAN (1968), tinha por objectivo a construção de um modelo, que designou de modelo Z, capaz de prever riscos de falência. Para o efeito, recorrendo à análise discriminante múltipla, propôs-se analisar como os rácios financeiros poderiam ser utilizados para identificar as empresas que entraram em falência no período de 1946 a 1965. Baseou-se, então, numa amostra de 66 empresas, com um volume de activos não rigorosamente idênticos e pertencentes ao mesmo sector de actividade, das quais 50% entraram em falência no período em referência. Os resultados obtidos pelo autor com a aplicação da metodologia foram os seguintes:

$$Z = 1,2 R_1 + 1,4 R_2 + 3,3 R_3 + 0,6 R_4 + 0,9 R_5$$

em que:

- Z : Índice sintético ou *score* discriminante;
- R<sub>1</sub> : Fundo de maneo sobre o activo total;
- R<sub>2</sub> : Resultados transitados retidos sobre o activo total;
- R<sub>3</sub> : Resultados antes de impostos e juros sobre o activo total;
- R<sub>4</sub> : Valor de mercado do capital próprio sobre valor contabilístico da dívida;
- R<sub>5</sub> : Vendas sobre o activo total.

Assim, o modelo através de R<sub>1</sub> entra em consideração com uma medida de solvabilidade; com R<sub>2</sub> toma uma medida de rendibilidade acumulada, que tem em conta a idade da empresa; com R<sub>3</sub> mede a rendibilidade dos activos independentemente do financiamento adoptado e impostos; com R<sub>4</sub> adopta uma medida de autonomia financeira; e, finalmente, com R<sub>5</sub> traduz a rotação dos capitais.

Parece importante destacar, por um lado, o facto do modelo ser constituído por um número reduzido de rácios que, por sua vez, são bastante utilizados na análise financeira das empresas e, por outro lado, a qualidade dos resultados obtidos. Para testar a capacidade discriminante das variáveis, o autor efectuou um teste F, à diferença entre as médias dos valores de cada grupo (quadro 1).

QUADRO 1

Médias dos rácios para as duas amostras e teste F de significância individual à diferença das médias

Rácios	Média nas empresas em falência	Média nas empresas Normais	teste F <sub>(1, 66)</sub>
R <sub>1</sub>	-6,1%	41,4%	32,60*
R <sub>2</sub>	-62,6%	35,5%	58,86*
R <sub>3</sub>	-31,8%	15,3%	26,56*
R <sub>4</sub>	40,1%	247,7%	33,26*
R <sub>5</sub>	150,0%	190,0%	2,84

Notas: \* quer representar teste significativo a um nível de 0,1% de significância.

Fonte: ALTMAN (1968).

Posteriormente, o autor procura determinar a contribuição relativa de cada variável para a capacidade discriminante total da função e sua interacção, tendo concluindo que as variáveis que contribuem mais para a separação dos grupos, são respectivamente R<sub>3</sub>, R<sub>5</sub>, R<sub>4</sub>, R<sub>2</sub> e R<sub>1</sub>.

O que é surpreendente é o facto do rácio R<sub>5</sub>, que era insignificante quando analisado isoladamente, passar a ter, quando combinado com os outros rácios,

um papel decisivo para a capacidade discriminante total da função<sup>1</sup>. Este resultado vem assim apontar para a superioridade da análise multidimensional face à unidimensional.

O quadro 2 resume os resultados obtidos. Constatamos que, nas previsões efectuadas com base no *score Z* um ano antes da falência de empresas, o modelo “acertava” em 94% dos casos para as empresas efectivamente falidas e, nas empresas efectivamente não falidas, os resultados melhoravam com o modelo a “acertar” em 97% dos casos. A precisão do modelo diminui substancialmente quando o número de anos aumenta, tendo o autor retirado duas importantes conclusões:

- que todos os rácios observados mostram uma tendência de deterioração à medida que se aproxima a falência;
- que as variações mais importantes da maioria dos rácios acontecem entre o terceiro e o segundo ano que precede a falência.

QUADRO 2

## Resumo dos resultados da Análise de ALTMAN (1968)

Situação efectiva	Previsões do Modelo			
	1 ano antes da falência		2 anos antes da falência	
	Falidas	Não falidas	Falidas	Não falidas
Falidas	94%	6%	72%	28%
Não falidas	3%	97%	6%	94%

Fonte: ALTMAN (1968).

A propósito dos resultados obtidos e tendo em conta as aplicações efectuadas, o autor concluiu que, as empresas com um *score Z* superior a 2,99 deveriam ser classificadas como “normais” e as empresas com um *score Z* inferior a 1,81 como empresas “em falência” (i.e., falida ou em tendência para a falência). O intervalo entre 1,81 e 2,99 corresponderia à “zona de ignorância”, de incerteza ou de indecisão. Posteriormente, procurando estabelecer um “valor de corte” para *Z* que contivesse um menor número possível de empresas mal classificadas, definiu o intervalo 2,67-2,68. Assim, uma empresa com um *score Z* superior a 2,68 seria uma empresa “em não falência”, enquanto que um valor inferior àquele indicaria que se trata de uma empresa “em falência”.

<sup>1</sup> Ou seja, em linguagem estatística pode dizer-se que é insignificante individualmente, mas é um dos rácios (o segundo) que mais contribui para a significância global da função discriminante. Para uma excelente discussão sobre esta problemática (de que uma variável não significativa isoladamente pode ser importante numa análise multidimensional) veja-se, por exemplo, COOLEY *et al.* (1962).

Apesar das versões actualizadas, aperfeiçoadas e, geralmente, secretas do modelo (original) Z serem utilizadas por várias instituições (nomeadamente pela banca), o modelo original continua presente e em uso. Citamos, no quadro 3, o caso de uma aplicação recente a uma amostra de grandes empresas que, face ao conhecimento que o mercado possui sobre essas empresas, evidencia que as empresas com uma pontuação de Z mais elevada são também as que possuem um risco inferior de incumprimento.

QUADRO 3

## Pontuações Z numa amostra de grandes empresas

Empresa	Pontuação Z
<i>Bethlehem Steel</i>	0,7
<i>Coca-Cola</i>	7,3
<i>Exxon</i>	3,2
<i>Hewlett-Packard</i>	3,8
<i>IBM</i>	1,1
<i>Merck</i>	4,3
<i>Occidental Petroleum</i>	0,7
<i>USAir</i>	0,6
<i>Wal-Mart</i>	6,2
<i>Westinghouse</i>	1,3

Fonte: BREALEY e MYERS (1998, pág. 856).

Na subsecção seguinte apresentamos um estudo de *ALTMAN et al.* (1977), que representou simultaneamente um avanço e uma síntese do estudo apresentado na presente subsecção.

### 2.1.2. O contributo de *ALTMAN, HALDEMAN e NARAYANAN (1977)*

Continuando a utilizar a análise discriminante na sua investigação, *ALTMAN et al.* (1977) chegaram a um modelo de sete variáveis, ao qual chamaram de modelo Zeta, que constitui uma evolução do modelo de cinco variáveis [de *ALTMAN (1968)*].

“*ALTMAN et al.* (1977) investigaram o desempenho de um modelo de análise com sete variáveis discriminantes (...). Em geral, o modelo das sete variáveis – denominado de modelo Zeta – é um aperfeiçoamento do modelo inicial de *ALTMAN* de cinco variáveis (*ALTMAN (1968)*.” [*ALTMAN et al.* (1998), pág. 1724].

Para construir o modelo Zeta os autores recorreram a uma amostra de 53 empresas falidas e seleccionaram previamente um conjunto de 27 rácios, sendo alguns deles uma inovação relativamente aos estudos anteriores, nomeadamente

ao apresentado na anterior subsecção. Algumas variáveis foram logaritmizadas de forma a reduzir a possibilidade de observações extremas e assim aderir com mais facilidade às hipóteses de normalidade. No final, e seguindo a metodologia do estudo de ALTMAN (1968), os autores propuseram então as seguintes sete variáveis explicativas:

- Z<sub>1</sub>: rendibilidade dos activos (resultado antes de juros e impostos sobre total de activos);
- Z<sub>2</sub>: estabilidade da rendibilidade (desvio padrão do resultado antes de juros e impostos sobre total de activos (tendo em conta os últimos dez anos));
- Z<sub>3</sub>: serviço da dívida [ $\log_{10}$  (resultado antes de juros e impostos sobre juros)];
- Z<sub>4</sub>: rendibilidade acumulada (resultados transitados e reservas sobre o total de activos);
- Z<sub>5</sub>: liquidez (activo circulante sobre as dívidas de curto prazo);
- Z<sub>6</sub>: capital social sobre capitais permanentes (média dos últimos cinco anos);
- Z<sub>7</sub>: dimensão  $\log_{10}$  do total de activos.

Relativamente aos valores estimados para os coeficientes associados às sete variáveis explicativas do modelo, os autores optaram por não os publicar atendendo a que servem de base a um serviço de informações privado, a *Zeta Services*. Porém, os autores não se abstiveram de dar algumas indicações, referindo que a variável menos importante é a rendibilidade dos activos e as variáveis mais importantes são a rendibilidade acumulada e a estabilidade da rendibilidade.

Neste estudo, os autores apresentam como grande avanço a referência à necessidade de se prever o resultado à partida (probabilidade *à priori*) e os diferentes custos, consoante o tipo de erro que se obtém na classificação das empresas: erro de 1ª espécie que consiste em classificar como empresa “boa” uma empresa “má”; e erro de 2ª espécie que consiste em classificar como empresa “má” uma empresa “boa”.

## 2.2. Estudo baseado em modelos econométricos do tipo *logit*: o contributo de KEASEY e MCGUINNESS (1990)

Um dos propósitos do trabalho de KEASEY e MCGUINNESS (1990) consistiu em determinar quais as variáveis financeiras mais relevantes para explicar a falência de empresas industriais do Reino Unido, no período 1976-1983. A amostra era constituída por 86 empresas das quais 43 eram falidas e 43 eram não falidas. Utilizando o modelo econométrico *logit*, os autores concluíram que os rácios mais importantes para o efeito, nos cinco anos imediatamente anteriores à falência, eram os de rendibilidade e de eficiência.

Refira-se que na selecção inicial do conjunto global de (potenciais) variáveis utilizadas (e constantes no quadro 4) para estimar as funções *logit*, KEASEY e MCGUINNESS (1990, pág. 121) tiveram em conta, por um lado, rácios frequentemente utilizados para analisar os principais aspectos do comportamento económico-financeiro das empresas e, por outro lado, os rácios utilizados por ZAVGREN (1985).

QUADRO 4

Conjunto global de (potenciais) variáveis

Variável dependente binária	$\left\{ \begin{array}{l} 1 = \text{Falência} \\ 0 = \text{Não – falência} \end{array} \right.$
Rácios financeiros	
Variáveis independentes (económico-financeiras):	Sinal esperado
1. RSE: rendibilidade dos capitais próprios(%)	-
2. RCE: rendibilidade do capital investido (%)	-
3. TPM: margem de lucro comercial(%)	-
4. PPM: margem de lucro antes de impostos(%)	-
5. TAE: volume de negócios sobre imobilizado corpóreo líquido	-
6. ST: rotação das existências (vendas sobre existências)	-
7. DT: rotação dos devedores (vendas sobre devedores)	-
8. CT: (vendas-lucro das vendas) sobre credores	-
9. CG: efeito alavanca do capital (%)	+
10. IG: efeito alavanca dos resultados(%)	+
11. BR: rácio do endividamento	+
12. WC: Grau de liquidez geral	-
13. QAR: Activo circulante sobre Activo	-
14. CA: disponibilidades sobre Activo	-
15. INS: 1/ST (existências sobre vendas)	+
16. RI: ST x (1/DT)	+

Fonte: KEASEY e MCGUINNESS (1990, pág. 121).

Finalmente, apresentamos no quadro 5 as funções *logit* estimadas pelos autores, “a melhor” para cada um dos cinco anos que antecede a falência, i.e., a que em cada um daqueles anos possui melhor poder discriminante de entre as funções estimadas e estatisticamente significativas. Conforme se pode constatar o número máximo de variáveis explicativas é de cinco e tal apenas acontece para a função relativa ao terceiro ano antes da falência.

A percentagem de classificações erradas diminui com a aproximação do ano da falência (1 ano: 14%; 2 anos: 18,5%; 3 anos: 23,5%; 4 anos: 30%; e, 5 anos 24,5%), como seria de esperar.

Por sua vez, no que diz respeito aos coeficientes associados às variáveis explicativas constatamos que os sinais estão de acordo com o esperado.

QUADRO 5

Função *logit* estimada para cada um dos cinco anos que antecede a falência

Anos	Função estimada para cada um dos anos	SG $\chi^2_{k,0.01}$	Classificação (%)		
			F <sup>a)</sup>	NF <sup>b)</sup>	Total <sup>c)</sup>
1	$P_f = 0,088 + 0,032CG - 0,271CT - 0,323PPM$ (0,081) (2,092)** (2,014)** (4,147)*	42,04*	86	86	86
2	$P_f = 3,36 + 8,43INS - 2,72WC - 0,108RCE - 0,020SER$ (2,931)* (2,882)* (3,364)* (3,093)* (1,605)	37,80*	84	79	81,5
3	$P_f = 6,420 - 1,560QAR - 0,300CT - 0,880TAE - 0,422PPM -$ $0,085RSE$ (3,769)* (1,519) (2,455)* (2,409)* (3,446)* (2,025)**	20,06*	72	81	76,5
4	$P_f = 0,498 + 0,035CG - 0,060CA - 0,204CT$ (0,644) (2,391)* (1,578) (2,250)**	27,89*	70	70	70
5	$P_f = 3,628 - 2,879QAR - 0,256CT$ (3,831)* (3,077)* (2,783)*	60,48*	70	81	75,5

**Notas:** - SG: Significância Global da regressão,  $\chi^2_{k,0.01}$  em que k corresponde ao número de variáveis independentes – graus de liberdade - e 0.01 ao nível de significância;  
- entre parênteses: estatística t;  
- \* e \*\* traduzem a rejeição da hipótese  $H_0$  [de ausência de significância individual dos coeficientes (estatística t) e de ausência de significância global da regressão] a 1% e 5%, respectivamente;  
- a), b) e c) percentagem de empresas falidas classificadas como falidas, percentagem de empresas não falidas classificadas como não falidas e percentagem de empresas totais classificadas correctamente, respectivamente.

Fonte: KEASEY e MCGUINNESS (1990, pág. 124).

### 3. ESTUDO EMPÍRICO

O objectivo desta secção é especificar e estimar um modelo tipo *scoring* (análise multidimensional) que com recurso a técnicas econométricas (utilizando o modelo *logit*) explique a probabilidade das empresas cumprirem as suas obrigações financeiras. Ou seja, propomo-nos modelizar e estimar uma função estocástica que, para características conhecidas de uma empresa, forneça a respectiva probabilidade de sobrevivência.

#### 3.1. A amostra e clarificação dos conceitos de empresa “má” e “boa”

Com base em 21 rácios económico-financeiros e nas taxas de variação de 8 obtidos através dos relatórios de contas tradicionais<sup>2</sup>, analisamos o seu comportamento para os dois conjuntos de empresas. Especificamente, analisa-se:

- um primeiro grupo de empresas que, segundo o sistema de informações do Banco Português do Atlântico (BPA), registaram incidentes não

<sup>2</sup> Os 21 rácios usados e as 8 taxas de variação foram fornecidos pela Central de Balanços do BPA, onde obtivemos a amostra.



justificados<sup>3</sup>, de 1991 a 1996, e designado pelo grupo das empresas “más”;

- um segundo grupo de empresas, escolhidas aleatoriamente a partir da totalidade das que foram tratadas pela Central de Balanços do BPA (CBBPA) e sem resposta de quaisquer incidentes, designado por empresas “boas”.

No caso de uma determinada empresa apresentar tais incidentes em dois ou mais anos, considerou-se empresa “má” no primeiro daqueles anos. Deste modo, dispondo-se das empresas que registaram tais incidentes em 1991, 1992, 1993, 1994, 1995 e 1996 foi possível constituir a amostra deste tipo de empresas. Para o efeito, tomando  $t+1$  como período de referência genérico<sup>4</sup>, obteve-se para cada empresa informação relativa aos momentos  $t$ ,  $t-1$  e  $t-2$ ,<sup>5</sup> com o objectivo de observar a divergência dos indicadores, em relação à média, um ano, dois anos e três antes da falência.

QUADRO 6

Empresas “más” em  $t+1$  e correspondentes períodos  $t$ ,  $t-1$  e  $t-2$ :

$T+1$	$T$	$t-1$	$t-2$
1991	1990	1989	1988
1992	1991	1990	1989
1993	1992	1991	1990
1994	1993	1992	1991
1995	1994	1993	1992
1996	1995	1994	1993
Nº observações	527	766	790

Um primeiro problema surgiu com a falta de sequência da informação ao longo do triénio, dentro do conjunto de empresas “más” (quadro 6). De  $t-2$  para  $t-1$  o número de empresas que entregaram as “contas” mantém-se praticamente constante, mas, pelo contrário, verifica-se que em  $t$  diminui consideravelmente; mais concretamente, cerca de 1/3 das empresas já não entrega “contas” no ano que antecede o surgimento do(s) incidente(s). Em rigor, este facto, limita então a previsão de incidentes, através de sinais de alerta ou outro qualquer processo, a  $t-1$  e  $t-2$ .

<sup>3</sup> Mais especificamente, apontes, cheques devolvidos, contencioso ou inibição de cheques no BPA ou no sistema bancário.

<sup>4</sup> Período em que a empresa foi classificada de “má” ou “boa”.

<sup>5</sup> Exemplificando: se, por exemplo, uma empresa foi classificada “má” em 1996,  $t$ ,  $t-1$  e  $t-2$  correspondem a 1995, 1994, 1993, mas as “contas” dessa empresa em cada um destes anos apenas são conhecidas e tratadas pelo Banco no ano seguinte (1996, 1995 e 1994 respectivamente).

Relativamente à selecção das empresas “boas” em cada período, três preocupações foram impostas: a primeira, quanto à forma de escolha, optou-se pela escolha aleatória dentro do conjunto de empresas com dados na CBBPA; a segunda, teve a ver com a necessidade de considerar cada empresa seleccionada apenas num período; a terceira, atendendo aos períodos amostrais desejados, levou-nos a encurtar o período temporal de escolha até 1994<sup>6</sup>.

## QUADRO 7

Empresas “boas” em  $t+1$  e correspondentes períodos  $t$ ,  $t-1$  e  $t-2$

$T+1$	$T$	$t-1$	$t-2$
1991	1990	1989	1988
1992	1991	1990	1989
1993	1992	1991	1990
1994	1993	1992	1991
Nº Observações	800	800	800

A escolha da variável “a empresa registou incidentes” enquanto indicador de situações de crise - correspondendo, eventualmente, ao último patamar de um percurso de crise - parece-nos perfeitamente justificada [veja-se, por exemplo, KOENIG (1985) e MQE (1995)], onde tal característica é típica de empresas cujo risco de falência é elevado]. Acresce que tal situação pode ser objectivamente avaliada garantindo a base sólida de análise que é necessária em qualquer modelo de previsão.

Não podemos terminar este ponto sem referir que, desde logo, existe uma limitação relativamente à ausência de dados por sector em cada região (pelo menos nas fontes a que tivemos acesso). Deste modo, ignoramos a dinâmica própria dos sectores nas regiões, dado que é possível admitir que os sectores e as regiões têm características próprias, não consideradas especificamente numa análise agregada. Porém, também não podemos esquecer que com a opção pelo módulo empresa resultou a possibilidade de construção de uma amostra de empresas “más” mais numerosa, que nos garante maior precisão nas estimativas obtidas para os coeficientes associados às variáveis explicativas.

<sup>6</sup> Data limite, porque pretendemos garantir que as empresas “boas” seleccionadas não registaram incidentes até ao momento da recolha de dados do BPA, ou seja, são realmente “boas” de acordo com a nossa definição. Com efeito, como o nosso objectivo é fazer projecções com os resultados alcançados, determinando-se, por exemplo, com base nas contas de  $t-2$ ,  $t-1$  e  $t$  uma determinada empresa entrará ou não em possível risco de crédito em  $t+1$ ; então, desde logo, nada nos garante que empresas sem incidentes, por exemplo, a partir de 1994 ( $t-2$ ) não venham a ter incidentes a partir de 1997 ( $t+1$ ).

### 3.2 O modelo proposto

O modelo que propomos é o *logit*, dado que é o indicado quando a variável explicada é dicotómica. Além disso, a este tipo de modelos estão associadas todas as vantagens da análise multidimensional e ainda algumas vantagens adicionais:

- os resultados de estimação dão-nos de uma forma simples a probabilidade de uma empresa ser uma “boa” empresa, isto é, obtemos um valor fácil de interpretar;
- proporciona, por via dos próprios valores que gera e teste que permite realizar, resultados convincentes.

Por outro lado, não estão sujeitos à principais críticas feitas à análise discriminante que, por exemplo, PINCHES (1980) e CHANG *et al.* (1999) apresentam.

Podemos, assim, começar por escrever a especificação linear geral:

$$Y_{i,t+1} = \alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k + u_{i,j} \quad (\text{III.1})$$

em que:

Y: variável *dummy* definida da seguinte forma:

$$Y_{i,t+1} = \begin{cases} 1 & \text{se a } i\text{-ésima empresa é "boa", no período } t+1; \\ 0 & \text{se a } i\text{-ésima empresa é "má", no período } t+1. \end{cases}$$

$X_{i,j}^1, \dots, X_{i,j}^k$ : variáveis explicativas, na  $i$ -ésima empresa, no período  $j$ ;

$\beta_1, \dots, \beta_k$ : parâmetros desconhecidos a estimar;

$\alpha$ : parâmetro constante desconhecido a estimar;

$u_{i,j}$ : termo de perturbação aleatório;

$j$ : período temporal que no nosso caso corresponde a  $t, t-1$  e  $t-2$ .

#### QUADRO 8

Valor assumido pela variável explicada e dimensão da amostra em cada um dos períodos que antecede a classificação

Variável	Designação	Dimensão da amostra		
		Em $t$	EM T-1	EM T-2
$Y_{i,t+1} = 1$	A $i$ -ésima empresa é “boa”, em $t+1$	800	<b>800</b>	800
$Y_{i,t+1} = 0$	A $i$ -ésima empresa é “má”, em $t+1$	527	<b>766</b>	790
	Total	1327	<b>1566</b>	1590

Donde, de acordo com a função distribuição *logit* e com base em (III.1) temos que:

$$P_{i,t+1} = P(Y_{i,t+1} = 1) = \frac{e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}}{1 + e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}} \quad (III.2)$$

pelo que:

$$P_{i,t+1} = P_{i,t+1} + u_{i,j} = \frac{e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}}{1 + e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}} + u_{i,j}.$$

Para estimar os coeficientes pelo método de máxima verosimilhança, assim como as estimativas dos respectivos desvios-padrões<sup>7</sup>, que permitem a análise estatística assintótica, a variável explicada terá de ser uma função linear desses coeficientes. A partir de (III.2) pode deduzir-se a especificação:

$$\ln\left(\frac{Y_{i,t+1}}{1 - Y_{i,t+1}}\right) = \alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k + u_{i,j}. \quad (III.3)$$

Com os resultados da estimação obtemos a probabilidade de uma empresa com características  $X_{i,j}^1, \dots, X_{i,j}^k$  ser uma “boa” empresa mediante a expressão (em que a acento circunflexo  $\hat{\phantom{x}}$  quer significar valor estimado):

$$\hat{P}_{i,t+1} = \frac{e^{\hat{\alpha} + \hat{\beta}_1 X_{i,j}^1 + \dots + \hat{\beta}_k X_{i,j}^k}}{1 + e^{\hat{\alpha} + \hat{\beta}_1 X_{i,j}^1 + \dots + \hat{\beta}_k X_{i,j}^k}}$$

Finalmente, antes de passarmos à estimação do modelo, note-se que há que ter algum cuidado na interpretação dos resultados do modelo *logit*. Enquanto os coeficientes estimados de acordo com a especificação linear (III.1) nos dão os efeitos marginais estimados das variáveis explicativas sobre a variável explicada, o mesmo não acontece com os coeficientes estimados por (III.3) já que aqueles resultam de:

<sup>7</sup> Os estimadores de máxima verosimilhança destes parâmetros têm distribuições assintóticas que, em rigor, só se verificam quando o número de observações tende para infinito. No entanto, em face de uma amostra obviamente finita (mas de dimensão significativa) e seguindo a prática corrente [veja-se, por exemplo, GREENE (1997)], assume-se que a amostra é “suficientemente grande”.

$$\frac{E(Y_{i,t+1})}{(X_{i,j}^1 + \dots + X_{i,j}^k)} = \frac{e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}}{1 + e^{\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k}} \times \frac{(\alpha + \beta_1 X_{i,j}^1 + \dots + \beta_k X_{i,j}^k)}{(X_{i,j}^1 + \dots + X_{i,j}^k)}$$

E, se os coeficientes estimados no modelo *logit* não admitem uma interpretação imediata, muito menos são comparáveis com os coeficientes estimados que se obtêm com o modelo linear de probabilidade. Por exemplo, AMEMIYA (1981) sugeriu que se fizesse a seguinte transformação para levar a cabo comparações entre as estimações:

$$\hat{\beta}_{\text{Modelo linear probabilidade}} = 0,25 \hat{\beta}_{\text{Modelo logit}}$$

$$\hat{\alpha}_{\text{Modelo linear probabilidade}} = 0,25 \hat{\alpha}_{\text{Modelo logit}} + 0,50$$

Apresentam-se de seguida os resultados da aplicação do modelo *logit* [especificação (III.3)] à amostra, efectuada com recurso ao programa de *software* econométrico *Limdep - version 7.0* [GREENE (1995)]. Ou seja, os quadros 9 [resultados de estimação da especificação (III.3), em  $t$ ,  $t-1$  e  $t-2$ ], 10 e 11 [que em conjunto nos dão ideia da qualidade dos ajustamentos, em  $t$ ,  $t-1$  e  $t-2$ ], resumem os resultados de estimação para  $t$ ,  $t-1$  e  $t-2$ .

As variáveis explicativas aplicadas em cada período resultaram da selecção dos indicadores com melhor desempenho na explicação do fenómeno em estudo em  $t$ ,  $t-1$  e  $t-2$ . Esta selecção foi efectuada através da construção de histogramas (que eliminou as taxas de variação por apresentarem elevado grau de sobreposição) e pela execução de testes estatísticos (testes  $t$  de Student, algumas medidas descritivas [de localização (média e mediana) e de dispersão (desvio-padrão e amplitude inter-quartil)]. Mais explicitamente, procuramos seleccionar o menor número de variáveis explicativas que garantisse a melhor significância global das regressões de acordo com, por um lado, a significância estatística individual dos coeficientes (mediante a observação da estatística  $t$  de *Student*) e, por outro lado, atender à análise unidimensional prévia em que procuramos os rácios de maior poder discriminante. Foram, assim, seleccionadas:

- para  $t$ : as variáveis explicativas encargos financeiros sobre os meios libertos totais<sup>8</sup> (EFMLT), encargos financeiros sobre a produção (EFPRD), encargos financeiros sobre o valor acrescentado bruto (EFVAB), autonomia

<sup>8</sup> Meios libertos = autofinanciamento (resultados líquidos + amortizações do exercício e variação de provisões - distribuição de dividendos) + encargos financeiros + imposto sobre o rendimento.

financeira (AUTON), produção sobre o ponto crítico (PRDPC) e o prazo médio de pagamento (PMP);

- para *t-1*: as variáveis explicativas EFMLT, EFPRD, AUTON, valor acrescentado bruto sobre a produção (VABPRD), PRDPC e PMP;
- para *t-2*: EFMLT, EFPRD, AUTON, resultados líquidos sobre os capitais próprios (RLCP), VABPRD, PRDPC, PMP e prazo médio de recebimentos (PMR).

QUADRO 9

Resultados de estimação da especificação (III.3), em *t*, *t-1* e *t-2*

Variáveis Explicativas	Coeficientes associados às variáveis explicativas, i.e., $\hat{\alpha}$ e $\hat{\beta}$ 's		
	Especificação <i>t</i>	Especificação <i>t-1</i>	Especificação <i>t-2</i>
Constante	0,0807	-1,1432	-1,4119
t-estatística <sup>b)</sup>	(0,2113)	(-3,5142)	(-3,9812)
Nível significância <sup>c)</sup>	[0,8334]	[0,0004]	[0,0001]
EFMLT	-0,0380	-0,0217	-0,0749
t-estatística <sup>b)</sup>	(-3,1531)	(-2,0624)	(-2,3714)
Nível significância <sup>c)</sup>	[0,0016]	[0,0623]	[0,0178]
EFPRD	-7,7528	-15,3790	-16,1290
t-estatística <sup>b)</sup>	(-3,9112)	(-9,9882)	(-9,6684)
Nível significância <sup>c)</sup>	[0,0001]	[0,0000]	[0,0000]
EFVAB	-1,7527		
t-estatística <sup>b)</sup>	(-3,4781)		
Nível significância <sup>c)</sup>	[0,0005]		
AUTON	2,3735	1,9294	2,1417
t-estatística <sup>b)</sup>	(5,1331)	(4,7692)	(5,4673)
Nível significância <sup>c)</sup>	[0,0000]	[0,0000]	[0,0000]
RLCP			-0,0629
t-estatística <sup>b)</sup>			(-3,5632)
Nível significância <sup>c)</sup>			[0,0004]
VABPRD		2,3508	1,5594
t-estatística <sup>b)</sup>		(5,0632)	(3,4512)
Nível significância <sup>c)</sup>		[0,0000]	[0,0006]
PRDPC	1,1560	0,3918	1,4340
t-estatística <sup>b)</sup>	(3,7412)	(4,9112)	(4,8902)
Nível significância <sup>c)</sup>	[0,0002]	[0,0000]	[0,0000]
PMP	-0,1183	-0,1376	-0,1009
t-estatística <sup>b)</sup>	(-3,8862)	(-5,0761)	(-3,1574)
Nível significância <sup>c)</sup>	[0,0001]	[0,0000]	[0,0015]
PMR			0,0416
t-estatística <sup>b)</sup>			(1,1774)
Nível significância <sup>c)</sup>			[0,2390]
Nº de observações	1327	1566	1590
Signif. Global ( $\chi^2$ ) <sup>a)</sup>	177,2933	180,7321	184,4752
Nível significância <sup>d)</sup>	[0,0000]	[0,0000]	[0,0000]

Nota: Os resultados estimados têm em conta o peso da amostra de cada grupo de empresas (isto é, "boas" e "más") na respectiva população.

QUADRO 10

Uma medida alternativa ao teste  $\chi^2$  para avaliar a qualidade dos ajustamentos

Classificação Efectiva das Empresas	Classificação de acordo com os resultados de estimação da especificação (III.3) <sup>r</sup>								
	Especificação t			Especificação t-1			Especificação t-2		
	0 <sup>p</sup>	1 <sup>q</sup>	Total	0 <sup>p</sup>	1 <sup>q</sup>	Total	0 <sup>p</sup>	1 <sup>q</sup>	Total
"má", i.e., 0	290	237	527	516	250	766	548	242	790
"boa", i.e., 1	89	711	800	189	611	800	209	591	800
Total	379	948	1327	705	861	1566	757	833	1590

QUADRO 11

Percentagem de classificações correctas e erradas

	t	t-1	t-2
Percentagem de erros	24,57%	28,03%	28,36%
Percentagem de classificações correctas	75,43%	71,97%	71,64%

Notas:

- Signif. Global ( $\chi^2$ ): teste de significância global com base no rácio de verosimilhança  $\chi^2$  correspondente à hipótese nula de os coeficientes associados às variáveis explicativas serem conjuntamente nulos;
- t-estatística: estimativas assintóticas da estatística t do estimador do parâmetro;
- Nível de significância: nível de significância a partir do qual se rejeita a hipótese nula de o coeficiente ser igual a zero, obtido na distribuição normal reduzida;
- Nível significância: nível de significância a partir do qual se rejeita a hipótese nula do teste de significância global com base em  $\chi^2$ ;
- 0: Classificação da empresa como "má", de acordo com os resultados de estimação da especificação (V.3);
- 1: Classificação da empresa como "boa", de acordo com os resultados de estimação da especificação (V.3);
- r) Classificação com base nos resultados de estimação da especificação (V.3), tomando como valor de corte a probabilidade de 0, 5;

Tomando os resultados estimados com base no modelo *logit* verificamos:

- através dos testes  $\chi^2$  do rácio de verosimilhança, que as três especificações, i.e., para t, t-1 e t-2 são globalmente significativas. Daí se conclui que as variáveis consideradas contribuem no seu conjunto, em cada caso t, t-1 e t-2, de forma significativa para a explicação da variável explicada;
- com base numa outra medida da qualidade global de estimação (constante dos quadros 10 e 11) que consiste na contagem do número de ocorrências amostrais que são errada e correctamente previstas pelos valores estimados, regista-se que a percentagem de classificações correctas é sempre superior a 71%, sugerindo uma boa qualidade dos ajustamentos. Mais especificamente, do quadro 10 conclui-se que:

– para t:

- das 527 empresas "más" a estimação "acerta" em 290 empresas classificando-as como tal, o que corresponde a um desempenho de 55,03%; o mau resultado deriva do facto já salientado de que cerca

- de 1/3 das empresas não entrega as contas no ano que antecede o primeiro incidente.
- das 800 empresas “boas” a estimação “acerta” em 711 empresas classificando-as como tal, correspondendo a um desempenho de 88,88%;
  - globalmente o modelo apresenta um desempenho de  $((290+711)/1327)$  75,43%;
- para  $t-1$ :
- das 766 empresas “más” a estimação “acerta” em 516 empresas, o que corresponde a um desempenho de 67,36%;
  - das 800 empresas “boas” a estimação “acerta” em 611 empresas classificando-as como tal, correspondendo a um desempenho de 76,38%;
  - globalmente o modelo apresenta um desempenho de 71,97%;
- para  $t-2$ :
- das 790 empresas “más” a estimação “acerta” em 548 empresas, o que corresponde a um desempenho de 69,37%;
  - das 800 empresas “boas” a estimação “acerta” em 591 empresas classificando-as como tal, correspondendo a um desempenho de 73,88%;
  - globalmente o modelo apresenta um desempenho de 71,64%;
- em qualquer dos casos ( $t$ ,  $t-1$  e  $t-2$ ) os erros são atribuíveis a variáveis explicativas relevantes (porventura de tipo qualitativo e, portanto, de difícil quantificação) não consideradas no modelo, a erros quanto à forma funcional da relação ou, caso exista coincidência absoluta entre o modelo estimado e a especificação correcta, à aleatoriedade da variável explicada. Naturalmente que a hipótese mais plausível é a de se ter verificado para a amostra utilizada um *mix* de todos estes efeitos; Além de que, dada a fonte de informação, há que ter sempre em conta a possibilidade do envio (por parte das empresas) do balanço melhorado para o BPA dada a finalidade a que se propõe (numa situação débil qualquer empresa com necessidade de recorrer ao crédito bancário terá tendência a retocar as suas contas dando uma imagem distorcida da sua real posição financeira) o que *per si* inviabiliza uma avaliação correcta das empresas;

quanto aos coeficientes estimados associados às variáveis individuais:



- à excepção do termo constante em  $t$  e da variável PMR em  $t-2$ , todas as restantes variáveis explicativas são estatisticamente significativas a níveis de significância desprezíveis (valores da estatística amostral  $t$  bastante significativos), demonstrando que cada uma das variáveis escolhidas é estatisticamente relevante para a explicação do fenómeno em causa;
- da análise dos efeitos marginais resulta que os sinais dos coeficientes estão conforme o esperado<sup>9</sup>.

Finalmente, de acordo com os resultados obtidos podemos então em  $t$ ,  $t-1$  e  $t-2$  classificar uma empresa de “boa” ou “má” em  $t+1$ . Assim, tomando, por exemplo, os resultados de estimação da especificação (III.3) para  $t-2$ , obteve-se:

$$\ln \frac{\hat{Y}_{i,t+1}}{(1 - \hat{Y}_{i,t+1})} = -1,4119 - 0,0749 \text{ EFMLT}_{i,t-2} - 16,1290 \text{ EFPRD}_{i,t-2} + 2,1417 \text{ AUTON}_{i,t-2} - 0,0629 \text{ RLCP}_{i,t-2} + 1,5594 \text{ VABPRD}_{i,t-2} + 1,4340 \text{ PRDPC}_{i,t-2} - 0,1009 \text{ PMP}_{i,t-2} + 0,0416 \text{ PMC}_{\text{RI},t-2} = \hat{Q}_{i,t-2}$$

donde,

$$\frac{\hat{Y}_{i,t+1}}{(1 - \hat{Y}_{i,t+1})} = e^{\hat{Q}_{i,t-2}} = \hat{R}_{i,t-2},$$

pelo que:

$$\hat{Y}_{i,t+1} = \frac{\hat{R}_{i,t-2}}{1 + \hat{R}_{i,t-2}} \quad (III.4)$$

O valor resultante de (III.4) corresponde à probabilidade estimada da  $i$ -ésima empresa ser “boa” (caso  $\hat{Y}_{i,t+1} > 0,5$ ) ou “má” (caso  $\hat{Y}_{i,t+1} < 0,5$ ) em  $t+1$  tendo em conta os valores das variáveis explicativas em  $t-2$ .

### 3.3. Exemplificação da aplicação dos resultados do Modelo *logit*, a uma empresa hipotética

Suponhamos que se pretendia saber a probabilidade de uma empresa hipotética registar incidentes em 1999, com base nos valores apresentados pelos rácios em 1997. Correspondendo 1999 ao período  $t+1$ , então 1997 corresponde a  $t-1$ , i.e., temos que utilizar a especificação estimada para  $t-1$ .

<sup>9</sup> À partida talvez fosse de esperar que o sinal associado à variável RLCP em  $t-2$  resultasse positivo; no entanto, o sinal negativo, sugerindo que à medida que esse rácio aumenta implica uma variação em sentido contrário na probabilidade da empresa registar incidentes, poderá ter a ver com o facto do seu aumento resultar mais de uma diminuição da CP do que de um aumento dos RL.

Admitamos, então, que o valor das variáveis explicativas na *i*-ésima empresa hipotética, em 1997 era de:

EFMLT: 0,24;  
 EFPRD: 0,03;  
 AUTON: 0,55;  
 VABPRD: 0,37;  
 PRDPC: 1,05;  
 PMP: 2,64.

Considerando os resultados de estimação da especificação (V.3) para *t-1*, obtém-se:

$$\begin{aligned} \ln \frac{\hat{Y}_{i,t+1}}{(1 - \hat{Y}_{i,t+1})} &= -1,1432 - 0,0217 \text{ EFMLT}_{i,t-1} - 15,3790 \text{ EFPRD}_{i,t-1} + 1,9294 \text{ AUTON}_{i,t-1} + \\ &+ 2,3508 \text{ VABPRD}_{i,t-1} + 1,3918 \text{ PRDPC}_{i,t-1} - 0,1376 \text{ PMP}_{i,t-1} = \\ &= -1,1432 - 0,0217 \times 0,24 - 15,3790 \times 0,03 + 1,9294 \times 0,55 + \\ &+ 2,3508 \times 0,37 + 1,3918 \times 1,05 - 0,1376 \times 2,64 = 1,419 \end{aligned}$$

donde,

$$\frac{\hat{Y}_{i,t+1}}{(1 - \hat{Y}_{i,t+1})} = e^{1,419} = 4,133$$

pelo que:

$$\hat{Y}_{i,t+1} = \frac{4,133}{1 + 4,133} = 0,805.$$

Assim, a probabilidade estimada da empresa ser “boa” em 1999, com base nos valores das variáveis explicativas de 1997 (i.e., com base nos seis rácios económico-financeiros que compõe a especificação (V.3) em *t-1*), é de 80,5% ou, dito de outro modo, a probabilidade da empresa ser “má” (i.e., da empresa vir a registar incidentes em 1999) é de 19,5%.

#### 4. CONCLUSÕES

Verificamos que o objectivo principal dos diferentes trabalhos de sinais de alerta de crise empresarial consiste em avaliar a situação das empresas através de algumas (poucas) variáveis tornando eventualmente desnecessária a análise tradicional mais demorada e custosa. Ou seja, procura-se proporcionar um trata-

mento preventivo, com base na detecção precoce facultada pelos modelos elaborados. Por sua vez, se as dificuldades são observadas quando já atingiram uma dimensão considerável, a análise explicativa permite, igualmente, identificar as causas e orientar de forma mais eficaz as medidas “curativas” necessárias para desencadear um plano de recuperação.

Com o presente trabalho procuramos então elaborar sinais de alerta de crise empresarial. Estamos conscientes que o fizemos de uma forma simples; desde logo porque apenas nos foi possível entrar em linha de conta com informação económico-financeira incorporada em rácios, obtida junto da CBBPA, não nos tendo sido possível recorrer a informação externa às empresas. Note-se, no entanto, que ainda assim os resultados obtidos merecem um olhar atento. Com efeito, os factores externos e internos “(...) exercem geralmente uma influência convergente no processo de degradação de uma empresa” [COHEN (1996, pág. 364)], pelo que os resultados conseguidos apenas com um daqueles dois tipos de dados não podem ser menosprezados. Refira-se, aliás, que estudos realizados no final dos anos 70 [por exemplo, CNME (1977) e CONAN *et al.* (1979)], anos 80 [por exemplo, CEPME (1986)] e anos 90 [DEPALLENS *et al.* (1990)] sugerem que a caminhada para a falência reflecte-se claramente na degradação progressiva da situação financeira da empresa. Assim, os principais estudos empíricos sobre o tema em análise tem, de facto, como principal fonte de informação a situação económico-financeira de uma empresa (i.e., informação interna), sobretudo expressa em rácios. Acontece assim com os estudos que constituem as principais referências: entre outros, BEAVER (1966), ALTMAN (1968), ALTMAN *et al.* (1977), COLLONGUES (1977), CONAN *et al.* (1979), VERDIER (1986), KEASEY *et al.* (1987) e KEASEY *et al.* (1990).

Constamos que esses estudos são oriundos de alguns (poucos) países (França, EUA e Inglaterra) e, em Portugal, quanto é do nosso conhecimento só MORGADO (1997), JULIANO (1998) e NEVES *et al.* (1998) fizeram contributos neste tema. Face ao exposto, o presente trabalho visa continuar a despertar o interesse do tema no nosso país.

Devemos, porém, ter em atenção que, por mais completo e, portanto, ausente de críticas que seja o estudo sobre previsão de falência das empresas, a sua aplicação não é válida para sempre, devendo ser actualizado permanentemente. Seja porque é de esperar que a realidade económico-financeira das empresas se vá alterando com o tempo, embora os estudos apresentem uma robustez temporal suficiente para que os factores conjunturais não alterem o seu significado, seja porque os desenvolvimentos ocorridos no tratamento de dados vão permitindo estimar modelos analíticos mais complexos.

Por outro lado, um estudo realizado sobre uma amostra de empresas num determinado país dificilmente poderá ser aplicado em outro país, pois a normalização contabilística à escala mundial não existe ainda, tendo ainda um longo

caminho a percorrer. Assim, o que fizemos tem a vantagem de atender à realidade empresarial portuguesa.

Finalmente, parece-nos importante deixar algumas reflexões para futuras investigações. Desde logo, para que os resultados obtidos sejam susceptíveis de dar garantias de qualidade seria interessante adoptar o tipo de análise efectuada a outras amostras. Poderíamos também comparar os resultados obtidos com o modelo *logit* com o modelo discriminante, contribuindo para a discussão sobre a eficiência dos modelos. Além disso, avançaremos com o desenvolvimento de outro tipo de modelos não estatísticos como, por exemplo, os modelos em lógica vaga (“fuzzy logic”) e os modelos baseados em redes neuronais.

**Referências:**

- ALTMAN, EDWARD I. - Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy -, *Journal of Finance*, Setembro de 1968, págs. 589-609.
- ALTMAN, E. I., HALDEMAN, R. G. e NARAYANAN, P. - Zeta Analysis: A new model to identify Bankruptcy risk of corporation -, *Journal Banking of Finance*, Junho de 1977, págs. 29-54.
- ALTMAN, E. I. e SAUNDERS, ANTHONY - Credit risk measurement: Developments over the last 20 years -, *Journal of Banking & Finance*, nº 21, 1998, págs. 1721-1742.
- AMEMIYA, T. - Qualitative response models: a survey -, *Journal of Economic Literature*, nº 19, 1981, págs. 483-536.
- BEAVER, W. H. - Financial Ratios as Predictors of Failure -, *Journal of Accounting Research Supplement, Empirical Research in Accounting: selected studies*, 1966, vol. 4, 1966, págs. 71-111.
- BREALEY, A. e MYERS, S. C. - *Princípios de Finanças Empresariais* -, MacGraw-Hill, 5ª Edição, 1998.
- CEPME (CRÉDIT D'EQUIPEMENT DES PME) - *Causes de défaillance des entreprises industrielles* -, Rapport, 1986.
- CHANG, LEE E SUNG - Dynamics of modeling in data mining: interpretative approach to bankruptcy prediction - *Journal of Management Informations Systems* - vol. 16, Summer 1999, págs. 63-85
- CNME (CAISSE NATIONALE DES MARCHÉS DE L'ÉTAT) - *Les causes de défaillance des entreprises commerciales* -, *Boletim nº 75*, 2º trimestre 1977.
- COHEN, ELIE - *Análise Financeira* -, Editora Presença, 1996.
- CONAN, JOEL e HOLDER, MICHEL - *Variables explicatives des performances et contrôle de gestion dans les PMI*, Thèse pour le Doctorat d'État, Université de Paris-Dauphine, 1979.
- COOLEY E LOHNES - *Multivariate Procedures for the Behavioral Sciences*, New York, John Wiley et Sons, Inc., 1962.
- DEPALLENS, G. e JOBARD J. P. - *Gestion Financière de l'entreprise* -, Editora Sirey, Paris - 10ª Edição - 1990.
- GREENE, WILLIAM - *Limdep version 7.0: user's manual* - Bellport, New York: Econometric Software, Inc., 1995.
- JULIANO, C. PIRES. *Modelos de Pontuação Quantitativa na Análise do Risco de Crédito*°. Dissertação de Mestrado. ISCTE. 1998.
- KEASEY, K. e MCGUINNESS, P. - The Failure of UK Industrial Firms for the Period 1976-1974, Logistic Analysis and Entropy Measures - *Journal of Business Finance and Accounting* - Primavera 1990, págs. 199-135.
- KEASEY, K. e WATSON, R. - Non-Financial Symptoms and the Prediction of Small Company Failure: A Test of Argenti's Hypotheses - *Journal of Business and Accounting*, *Autum*, 1987, págs. 335-354.
- KOENIG, GÉRARD - La vulnérabilité des entreprises -, *Révue de l'ISMEA*, 1985.
- MINISTÉRIO PARA A QUALIFICAÇÃO E O EMPREGO (MQE) - *Sistema de Indicadores de Alerta* -, Coleção Estatísticas, Departamento de Estatística, Outubro 1995.
- MORGADO, A. VILARES. *Análise Z-Score: Estudo teórico e prático no âmbito do enquadramento e características das empresas portuguesas*. Dissertação de mestrado. ISCTE. 1997.
- NEVES, J. CARVALHO E J. ANDRADE E SILVA. *Modelos do Risco de Incumprimento para a Segurança Social*. Fundação para a Ciência e Tecnologia/Instituto de Gestão Financeira da Segurança Social. 1998.
- PINCHES, GEORGE - Factors influencing classification results from multiple discriminate analysis - *Journal of Business Research*, 1980, págs. 429-456.
- VERDIER, MICHEL - Laide a Lanalyse Financière: Un Système de Prévention des Difficultés des Entreprises - *La Révue Française de Comptabilité*, nº 170, Julho-Agosto 1986, págs. 15-21.

**Abstract**

Various models have been developed to predict corporate failure. The earliest proposals correspond to univariate type model where a single variable, such as an accounting ratio, is examined and compared against the industry average ratio. Such an approach is too simple because it doesn't take into account the fact that ratios are not distributed normally around the average. Furthermore, this approach doesn't allow for interactions between variables. As a consequence of these shortcomings, over the last three decades multivariate failure prediction models have been developed. These new models consider the simultaneous interaction between variables, such as profitability, gearing, and liquidity. A first procedure widely used was the multiple discriminant analysis. More recently, it has become commonplace the use of logit or probit regression models because they have the potential advantage of being non-linear.

This paper is divided in two major parts: in a first part we survey some of the relevant literature; in a second part we present a logit regression model to predict corporate failure in Portugal.

Key words: Failure prediction models; univariate and multivariate models; discriminant analysis; logit regression model; Portuguese case.

---