

Previsão de insolvência de micro e pequenas empresas aplicando a Análise Discriminante

Elisa Henning (Universidade do Estado de Santa Catarina) elisa.henning@udesc.br
Fernando Natal de Pretto (Universidade do Estado de Santa Catarina) dep2fnp@joinville.udesc.br
Bruno Alexandre Pasold (Universidade do Estado de Santa Catarina) bapasold@yahoo.com.br

Resumo:

Para que uma instituição bancária possa sobreviver em meio a tantos riscos que a cercam, faz-se necessário o desenvolvimento e aplicação de métodos que sejam aptos a prever situações de potenciais perdas financeiras. Neste contexto encontram-se métodos estatísticos multivariados, como a análise discriminante. Assim, o presente trabalho traz a aplicação da análise discriminante com objetivo de realizar uma previsão quanto à adimplência ou inadimplência de micro e pequenas empresas junto a uma instituição bancária. Foi desenvolvido um modelo que utiliza informações já disponíveis no próprio sistema corporativo da instituição e outras obtidas com o próprio cliente. A amostra de análise foi composta de 58 empresas, 28 consideradas solventes e 28 insolventes, com 56 variáveis. O modelo resultante compõe-se de 11 variáveis, relacionadas à capacidade de cumprimento de obrigações financeiras junto a clientes e fornecedores, tanto da empresa quanto dos sócios. Através do presente estudo, foi possível verificar que essas informações já são úteis para a realização de uma análise quanto à solvência da empresa. O modelo desenvolvido, classificando as empresas em solventes e insolventes, apresentou uma taxa de erro global de 17,86% na amostra de análise. Foi delimitada também uma região crítica (de penumbra), com duas linhas de corte, em que a solvência da empresa não pode ser prevista. Para este caso a taxa de erro global foi de 8,9%.

Palavras chave: Análise Discriminante, Previsão de insolvência, Micro e pequenas empresas.

Predicting Bankruptcy in Small and Medium-Sized Businesses Using Discriminant Analysis

Abstract

For a bank to survive with so many risks that surround it, it is necessary to develop and apply methods that are able to predict situations of potential financial losses. In this context, there are multivariate statistical methods, such as discriminant analysis, that can be used. Thus, this paper presents the application of discriminant analysis in order to make a prediction regarding the repayment or default of loans to small and medium-sized businesses from a bank. We developed a model that uses information already available in the bank's own corporate system, and other information obtained from the customer. The analysis sample was composed of 58 companies, 28 considered solvent and 28 insolvent, with 56 variables. The resulting model consists of 11 variables related to the ability to meet financial obligations to customers and suppliers, both of the business and the shareholders. Through this study, we found that this information is already useful for performing an analysis regarding the solvency of the company. The developed model, classifying the companies into solvent and insolvent, presented an overall error rate of 17.86% for the analysis sample. A critical region (overlap area) was also delineated, with two cut-off lines, within which the company's solvency cannot be predicted. In this case, the overall error rate was 8.9%.

Key-words: Discriminant Analysis, predicting insolvency, small and medium-sized businesses.

1. Introdução

Para que uma empresa possa crescer e se desenvolver, são necessários vários ingredientes que contribuam para a consecução destes objetivos. Um deles é a disponibilidade de capital de giro e/ou de capital para investimentos, os quais podem estar disponíveis tanto com recursos próprios quanto através de crédito junto a terceiros. Como muitas empresas não são sólidas no mercado e, por consequência, não possuem disponibilidade financeira que subsidie todos os seus gastos com capital de giro e com novos investimentos, insere-se a importância da atividade bancária na economia e na sociedade de uma determinada região, cidade, estado e/ou país, mediante o Crédito que é disponibilizado àqueles que não possuem recursos suficientes.

Do ponto de vista econômico, a importância do crédito para empresas inicia-se com o suprimento da necessidade de capital de giro, as quais não dispõem dos recursos necessários para compra de matéria-prima e outros insumos momentaneamente, mas que precisam produzir para poder vender a seus clientes. O crédito também pode ser utilizado em modalidades de investimentos, ou seja, através de financiamento de máquinas, implementos ou qualquer outro empreendimento que aumentará as receitas da empresa e, muito provavelmente, sua lucratividade. Além dos benefícios que são trazidos diretamente à empresa recebedora do crédito, muitas outras vantagens econômicas e sociais ocorrem na região a qual está inserida, como por exemplo: a empresa demandante de crédito mitiga seu risco de falência por descasamento de caixa; empresas fornecedoras produzirão mais para atender à demanda de seu cliente; a economia regional é aquecida, pelo fato dos fornecedores terem que contratar novos funcionários para atender a demanda e, com o aumento dos funcionários, mais pessoas poderão fazer compras, pois possuem recursos para isso; em algum momento da cadeia de suprimentos, possivelmente as empresas clientes serão fornecedoras dos atuais fornecedores da empresa demandante do crédito, quando haverá aumento da demanda dos produtos a serem produzidos pela empresa em questão; com o aumento de funcionários na região, incrementar-se-á a demanda de produtos pelos consumidores finais, já que mais pessoas possuirão poder de compra, aumentando assim a produção de todas as empresas da região.

Para que todos estes benefícios para a economia e sociedade sejam conseguidos, é necessário que o processo de análise de crédito seja feito de forma correta e adequada para cada cliente ou para cada tipo de cliente. A insolvência se caracteriza pela incapacidade de empresas em cumprir suas obrigações na medida em que os vencimentos ocorrem (SILVA *et al.*, 2012). E, para Pinheiro *et al.* (2007) a insolvência é uma das dificuldades às quais as empresas são suscetíveis. Os sucessos e fracassos empresariais são inerentes ao capitalismo, todavia seria recomendável a previsão antecipada do fracasso de uma empresa, evitando assim consequências deste fato às organizações interessadas (SILVA *et al.*, 2012).

Baseado nisso, foi desenvolvido um modelo, a partir da Análise Discriminante, que consiste na previsão, embasada em dados históricos, da solvência (ou não) das micro e pequenas empresas clientes de uma instituição bancária. Entenda-se como micro e pequenas empresas aquelas que possuem Receita Operacional Bruta Anual de até R\$ 16 milhões, de acordo com a classificação do BNDES - Banco Nacional de Desenvolvimento Econômico e Social.

O desenvolvimento deste trabalho justifica-se pelo fato dos modelos de análise discriminante aplicada à análise de solvência de empresas, até hoje encontrados na literatura publicada, utilizarem informações de relatórios gerados somente para empresas de maior porte, como Balanço Patrimonial - BP e Demonstração de Resultados do Exercício - DRE. Por isso, buscou-se a elaboração de um modelo que utilizasse informações disponíveis no próprio

sistema corporativo da Instituição Bancária e outras extraídas do próprio cliente, do ramo de atividade em que atua e do conceito que o mesmo recebe tanto da própria instituição quanto da praça em que atua.

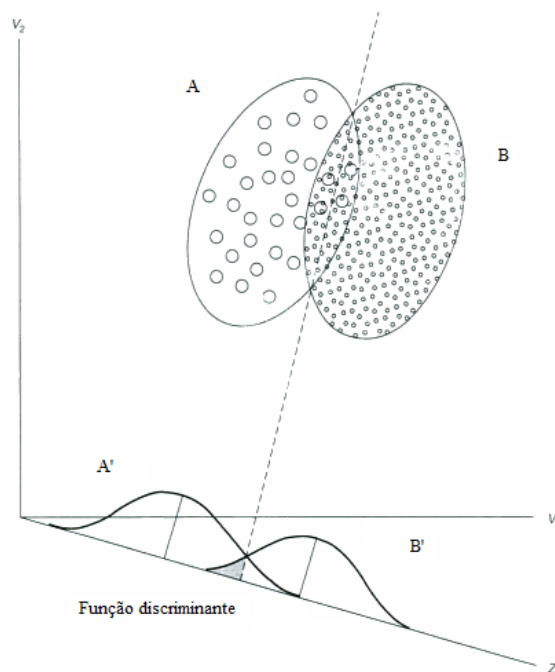
Assim sendo, a partir da presente introdução este artigo está delineado da seguinte forma: na Seção 2 apresenta-se o referencial teórico, descrevendo-se brevemente a análise discriminante e trabalhos correlatos; na Seção 3 estão os procedimentos metodológicos adotados e descrição da amostra; na Seção 4 apresentam-se os resultados e análise dos mesmos e, finalizando na Seção 5 estão as conclusões e considerações finais.

2. Referencial teórico

2.1 Análise Discriminante

No contexto da estatística multivariada, um dos principais questionamentos se refere a aplicação de suas técnicas para fins de classificação e posterior previsão dos elementos que estão sendo analisados. Problemas de classificação são comuns no cotidiano e a análise discriminante é uma das técnicas multivariadas que foram desenvolvidas com o propósito de auxiliar na classificação de um elemento em determinado grupo (MARIO, 2012).

A análise discriminante é, portanto uma técnica estatística que auxilia a identificar quais variáveis que diferenciam os grupos e quantas dessas variáveis são necessárias para se obter uma melhor classificação dos indivíduos de uma determinada população. A característica básica desta técnica é a utilização de um conjunto de informações obtidas acerca das variáveis independentes para conseguir um valor de uma variável dependente com a finalidade de classificá-la. A natureza da variável dependente é qualitativa (não métrica) podendo ser categórica, com uma classificação pré-estabelecida como, por exemplo, aprovado ou reprovado (MARIO, 2012; HAIR *et al.*, 2005). A Figura 1 apresenta uma ilustração gráfica da Análise discriminante de dois grupos.



Fonte: Adaptado de Hair *et al.* (2005).

Figura 1 – Ilustração gráfica da análise discriminante de dois grupos.

Com relação às variáveis independentes, estas geralmente são métricas, com valores contínuos, mas também podem assumir valores que representem categorias. Além da principal aplicação, que são as classificações dicotômicas a análise discriminante também pode ser utilizada para mais categorias. Deste modo um modelo de risco de crédito pode ser dividido em categorias de clientes rotuladas como de baixo, médio e alto risco, possibilitando gerar informações mais relevantes para os gestores. Em resumo, a Análise Discriminante auxilia a identificar se existem diferenças significativas entre as variáveis de cada grupo; reconhecer as variáveis que melhor diferenciam os grupos de observações; descrever uma ou mais funções que representem as diferenças entre os grupos e classificar a priori novos indivíduos nos grupos, com base na função discriminante aproximada (MARIO, 2012).

A função linear discriminante aplicada nesta pesquisa assemelha-se a uma equação de regressão múltipla, sendo composta pelas variáveis independentes que representam as características dos elementos, ponderadas de acordo com a sua importância ou impacto que causam na variável dependente. Tem como base a derivação de uma combinação linear das variáveis independentes que vão melhor discriminar entre os grupos pré-definidos. Essa discriminação é atingida com a atribuição de pesos para as variáveis, de forma a maximizar a variância entre os grupos em relação à variância dentro dos grupos. Pode ser resumida pela Equação (1)

$$Z = a + b_1X_1 + b_2X_2 + \dots + b_nX_n, \quad (1)$$

com Z correspondendo à variável dependente categórica, indicando uma pontuação ou escore discriminante; a é o intercepto da função quando $X_t = 0$; b_n é o coeficiente discriminante (o peso de cada variável na função) e X_n são os valores das variáveis independentes.

De modo geral, é recomendável que se tenha dados suficientes para duas amostras. A primeira amostra é aplicada para a estimação dos coeficientes, obtendo a função discriminante e análise das relações entre as variáveis preditoras e o problema de classificação. Já, a segunda amostra, de validação, é usada para avaliação final da capacidade classificatória do modelo. Embora a teoria recomende a divisão nestas duas amostras, há instrumentos de validação para o caso de amostras pequenas, quando não é possível esta divisão. A técnica de validação cruzada, é uma alternativa, em que uma função é gerada a partir de $n-1$ elementos de toda a amostra para se testar se esses elementos são corretamente classificados (MARIO, 2012).

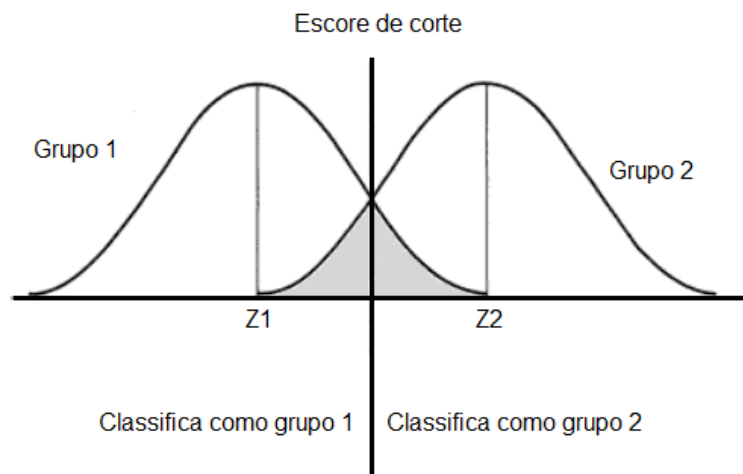
Os pressupostos da Análise Discriminante abrangem a normalidade multivariada, linearidade, ausência de “outliers”, ausência de multicolinearidade e homogeneidade das matrizes de variância-covariância. A suposição de normalidade multivariada pressupõe que a combinação linear entre as variáveis das funções segue uma distribuição normal. Sua ausência pode causar problemas de estimação na função discriminante. É recomendável, portanto, a avaliação da normalidade multivariada. A não observação da linearidade pode reduzir a robustez dos testes. A presença de “outliers” prejudica fortemente a análise discriminante e presença de multicolinearidade entre as variáveis incorre na “distorção” do efeito dos coeficientes. A última premissa, da homogeneidade das matrizes de variância e covariância tem como objetivo evitar que a função discriminante classifique equivocadamente observações em grupos de maior variância (MARIO, 2012; HAIR ET AL, 2005).

A partir desta função discriminante, parte-se para a construção da regra de classificação, por meio do escore Z discriminante. Este escore representa o resultado composto da multiplicação

de cada variável independente por seu peso correspondente. A discriminação dos grupos se dá através do estabelecimento dos pesos da variável estatística para cada variável regressora, com objetivo de maximizar a variância entre grupos relativa à variância dentro dos grupos. Deste modo, pode-se calcular o(s) centroide(s), que é o cálculo da média dos escores discriminantes para todos os indivíduos ou objetos. Para grupos de tamanhos iguais (o caso deste trabalho) os escores Z podem ser calculados mediante a expressão (2)

$$Z = \frac{Z_1 + Z_2}{2}, \quad (2)$$

onde Z corresponde escore de corte crítico para amostras de mesmo tamanho; Z_1 = centroide do grupo 1 e Z_2 o centroide do grupo 2. Este valor de Z representa o ponto de corte dos grupos. Valores de escores superiores a Z implicam que a observação faz parte do grupo 2 e inferiores a Z significam a observação é classificada como do grupo 1. Essa situação pode ser visualizada Figura 1 (HAIR *et al.*, 2005; MATEUS, 2010)



Fonte: Adaptado de Hair *et al.* (2005).

Figura 1: Escore de corte ótimo com tamanhos iguais de amostra.

Outra abordagem diz respeito à análise dos indivíduos ou objetos, pertencentes aos dois grupos, inseridos na intersecção das curvas de distribuição (Figura 1). Esta intersecção ou superposição das curvas denomina-se nome de “Zona de Indecisão” (ou também de penumbra) pois representa uma região de dúvida (*overlap area*) na classificação dos elementos distribuição das amostras dos dois grupos (MARIO, 2012; MATEUS 2010).

A análise da zona de indecisão permite avaliar se o modelo discriminante de dois grupos satisfaz à investigação pretendida ou se é necessário a utilização de um modelo múltiplo, ou seja, mais de dois grupos discriminantes. Os limites desta região podem ser definidos calculando-se um intervalo de confiança, de acordo com as Equações 3 e 4 (MARIO, 2012; MATEUS 2010).

$$X_1 = \bar{z}_2 - \sqrt{1 + \frac{1}{n_2}} t_{\alpha;gl} \sigma, \quad (3)$$

$$X_2 = \bar{z}_1 - \sqrt{1 + \frac{1}{n_1}} t_{\alpha;gl} \sigma, \quad (4)$$

onde \bar{z}_1 e \bar{z}_2 são as médias dos valores de Z calculados para as amostras de cada grupo; n_1 e n_2 correspondem a quantidade de elementos em cada amostra dos grupos; $t_{\alpha,gl}$ denota a estatística t de Student para o nível de confiança desejado e graus de liberdade; σ é o desvio padrão obtido a partir da variância ponderada dos grupos.

2.2 Trabalhos correlatos

O problema da ocorrência de crises bancárias tem merecido estudos desde a década de 30, quando foi realizado o primeiro trabalho no sentido de se prever a falência de bancos (SECRIST, 1938; PANDELÓ JR; PANDELÓ, 2008). A aplicação da análise discriminante relacionado ao diagnóstico do desempenho de empresas aparece no final da década 60, quando um dos artigos mais relevantes, de acordo com Mario (2012), foi o de Altman em 1968, inaugurando-se uma série de estudos sobre a capacidade preditiva de índices extraídos de demonstrações financeiras. Este assunto, a modelagem da insolvência bancária foi então retomado, de forma mais intensa, na década de 70, sendo que ALTMAN (1976, 1977, 1978) e SINKEY (1975), entre outros, desenvolveram estudos com o objetivo de antecipar a insolvência de associações de poupança e empréstimo ou bancos comerciais (PANDELÓ JR; PANDELÓ, 2008). Outro avanço, ainda segundo Mario (2012), foi o surgimento de modelos de ranqueamento de crédito (Credit Scoring Models) utilizando a análise discriminante para classificação de clientes no momento da concessão de crédito, assim como seu acompanhamento posterior.

No Brasil, o primeiro modelo de previsão de insolvência surgiu em 1976 (ELIZABETZKY, 1976; CORREIA, COSTA, MATIAS, 2006) sendo que diversos estudos dessa natureza foram realizados. Várias técnicas foram aplicadas, entre elas a Análise de Índices de Risco, Análise de Sobrevivência, Regressão Logística, Redes Neurais e a Análise Discriminante, cada qual com suas especificidades. No entanto, a maior parte dos modelos de previsão de solvência faz uso da análise discriminante como ferramenta estatística (SILVA *et al*, 2012). Uma limitação dos modelos, para Pandeló Jr e Pandeló (2008) é o número reduzido de amostras nos trabalhos nacionais, quando comparados com similares no exterior.

Deste forma, de acordo com exposto anteriormente, a análise discriminante configura-se uma boa opção para aproximação do modelo proposto.

3 Procedimentos metodológicos

O trabalho foi desenvolvido como um estudo empírico, caracterizando-se uma pesquisa exploratória. O objetivo é analisar a solvência de micro e pequenas empresas clientes de uma instituição bancária, de uma agência localizada na região norte do estado de Santa Catarina. A amostra constituiu-se inicialmente de 28 clientes classificados como “solventes” e 28 como “insolventes” sendo levantados dados cadastrais já presentes no sistema corporativo da instituição. Foi considerada como “solvente” aquela empresa que paga suas obrigações pontualmente, com raras exceções de atraso, mas não significativas, ou seja, num período inferior a cinco dias do vencimento. Como insolventes foram consideradas as empresas que interromperam o pagamento de suas dívidas ou que com atrasos superiores à trinta dias. Empresas que pagam com frequentes atrasos inferiores há trinta dias foram desconsideradas por estarem na época desta pesquisa numa área de “penumbra”, não podendo ser classificadas como solventes ou insolventes.

As informações coletadas da empresa, à época da pesquisa, foram as seguintes: ramo de atividade, data de início da operação, Capital Social, faturamento bruto anual, valor dos bens e patrimônio, tipo de propriedade do imóvel da empresa, a pontualidade da empresa junto a fontes externas, a pontualidade junto à própria Instituição, se é ou não cliente de outros

bancos, percentual de concentração de vendas nos cinco principais clientes, se possui ou não seguro empresarial, se possui ou não crédito junto a bancos e fornecedores, se possui ou não dificuldade pagar débitos de funcionamento, percentual da margem de lucro sobre as vendas, quantidade de cheques devolvidos nos últimos doze e seis meses, quantidade de dias de conta estourada nos últimos doze meses, quantidade de funcionários, valor dos recursos computáveis, valor e prazo das obrigações junto ao SFN (Sistema Financeiro Nacional), tempo de conta, total de compras efetuadas nos últimos três meses, utilização média do cheque especial nos últimos doze e seis meses, tipo de restrição (se existente) nos últimos cinco anos, valor médio da folha de pagamento nos últimos três meses; quanto aos sócios: se bem conceituado(s) na praça, se pontual(is) junto a fontes externas, tempo de experiência no ramo de atividade, dedicação do(s) sócio(s) às atividades da empresa (integral ou parcial), quantidade de cheques devolvidos nos últimos doze meses, quantidade de dias com a conta estourada nos últimos doze meses, quantidade de restrições nos últimos três anos, valor dos recursos computáveis, tempo de conta, utilização média do cheque especial nos últimos doze meses e o percentual de participação de cada sócio na empresa. Todas estas informações foram disponibilizadas pela Instituição Bancária, sem, porém se identificar de qual empresa os dados se referiam. Inicialmente o banco de dados apresentava 58 variáveis para cada cliente.

Em seguida foram excluídas as variáveis que não influenciariam na concessão de crédito aos clientes analisados, ou seja, todas as empresas forneceram às mesmas respostas frente estas variáveis. Foi, neste caso, feita a transformação de variáveis qualitativas em quantitativas por meio de utilização de variáveis *dummy*. E, neste momento foi também definida a variável *dummy* “adimplência”. Para esta variável o número “0” (zero) corresponde a inadimplência (empresas insolventes) e o “1” (um) a adimplência (empresas solventes).

A análise foi feita com auxílio do ambiente R (R Core Team, 2013) utilizando-se o pacote DiscrMiner (SANCHEZ, 2013) e o pacote MASS (VENABLES; RIPLEY, 2002). O R é uma linguagem e um ambiente para computação estatística, baseado no conceito de software livre, sendo multi-plataforma, contendo versões para Windows, MacOS, GNU/Linux e Unix). A função discriminante linear foi aproximada com auxílio dos pacotes DiscrMiner (SANCHEZ, 2013) e MASS (VENABLES; RIPLEY, 2002), assim como foram utilizadas funções destes pacotes para posterior classificação e validação do modelo. O pacote DiscrMiner contém funções para Análise Discriminante para fins de classificação abrangendo vários métodos, como análises descritivas, linear, geométrica, quadrática, assim como qualitativa. Para análise da normalidade multivariada foi utilizado o pacote QRM (PFAFF; McNEIL, 2012).

A partir de matriz de correlações foram excluídas também variáveis que estavam diretamente relacionadas, em razão da multicolinearidade. Foi avaliada a exclusão de *outliers* quando da análise de algumas variáveis quantitativas, os quais poderiam afetar os resultados da função discriminante. Avaliou-se também a normalidade multivariada a partir de um teste de normalidade multivariado baseado na distância de Mahalanobis, podendo-se aceitar a suposição de normalidade multivariada ($\alpha = 1\%$).

4. Resultados e análises

A partir da avaliação dos pressupostos, a amostra final resultou em 56 empresas, 28 solventes e 28 insolventes. Das 58 variáveis originais (Anexo A), uma primeira seleção de variáveis realizada por Pasold (2010), resultou num conjunto de 38 variáveis. Neste trabalho, partiu-se destas 38 variáveis selecionadas por Pasold (2010). Para a realização da análise discriminante o tamanho da amostra é fundamental e, de acordo com Hair *et al.* (2005) o ideal é que a razão entre o tamanho da amostra e quantidade de variáveis regressoras seja de 20 (vinte), com pelo menos 5 (cinco) amostras por variável regressora. Assim, sendo foram escolhidas as variáveis

que apresentavam maior correlação com a variável dependente. O modelo final resultante ficou com 11 variáveis explicativas.

Inicialmente apresentam-se na Tabela 1 os coeficientes da função discriminante aproximada e algumas medidas referentes à análise discriminante realizada. O coeficiente λ de Wilks e a estatística F são usados para avaliar a significância das variáveis. A partir destes resultados, verifica-se o que a variável mais significativa é a x_{21} (Quantidade de dias com excesso na conta corrente ou cheque especial nos últimos 12 meses) seguida da x_{31} (Valor das restrições SERASA últimos 5 anos – impeditiva). Poderia-se também avaliar a retirada das variáveis x_{19} , x_{20} , x_{24} , x_{44} e x_{56} , uma vez que o nível de significância é superior a 0,10.

	Variável	Coefficiente	λ Wilks	F	p-valor
x_{21}	Quantidade de dias com excesso na conta corrente ou cheque especial nos últimos 12 meses	-0,015690	0,8144	12,3	0,0009
x_{31}	Valor das restrições SERASA últimos cinco anos - impeditivas	-0,042420	0,9217	4,58	0,0368
x_{49}	Sócio 2 - pontualidade fontes externas - bem conceituado	-0,978700	0,9303	4,05	0,0493
x_{41}	Sócio - Quantidade restrições últimos três anos	-0,010620	0,9407	3,40	0,0705
x_{11}	Pontualidade nos pagamentos junto à Instituição - pequenos atrasos (até 30 dias)	-0,570100	0,9434	3,24	0,0774
x_{17}	Possui seguro empresarial - sim, na própria instituição	1,512000	0,9495	2,87	0,0959
x_{20}	Quantidade de cheques devolvidos nos últimos seis meses	-0,066970	0,9540	2,60	0,1124
x_{56}	Sócio 2 - tempo de conta (anos)	0,017200	0,9550	2,55	0,1163
x_{44}	Sócio - Utilização média do cheque especial últimos 12 meses	-1,162000	0,9574	2,40	0,1269
x_{24}	SFN - Endividamento a vencer em até 360 dias	-0,000003	0,9575	2,30	0,1273
x_{19}	Quantidade de cheques devolvidos nos últimos 12 meses	-0,010820	0,9609	2,19	0,1443

Fonte: Os autores

Tabela 1 – Coeficientes e medidas (λ de Wilks, F) da análise discriminante

Percebe-se que as variáveis significativas estão relacionadas à capacidade de cumprimento de obrigações financeiras junto a clientes e fornecedores, tanto da empresa quanto dos sócios.

A Equação 5 sumariza a expressão resultante para o modelo.

$$y = 0,290 - 0,5701x_{11} + 1,512x_{17} - 0,01082x_{19} - 0,066970x_{20} - 0,015690x_{21} - 0,000003x_{24} - 0,42420x_{31} - 1,162x_{44} - 0,9787x_{49} + 0,0172x_{56} \quad (5)$$

Na Tabela 2 estão os escores e a classificação das empresas da amostra, considerando a função discriminante ajustada. O ponto de corte é $Z = 0$ e, deste modo, empresas com escores positivos são classificadas como adimplentes (solventes) e com escores negativos como inadimplentes.

Em sequência (Tabela 3) estão os totais de valores observados e preditos pelo modelo (Equação 5). A taxa de erro global é de 17,86%, ou seja, o modelo acerta cerca 82% dos casos. Analisando os resultados da Tabela 3, verifica-se que a quantidade de acertos do “0” (insolvência) é maior do que as do “1” (solvência).

Empresa	Score (z)	Classificação	Observado	Empresa	Score (z)	Classificação	Observado
22	-3,53274	0	0	38	0,15938	1	1
19	-3,36844	0	0	25	0,18860	1	0
16	-2,06326	0	0	30	0,19965	1	1
12	-2,04232	0	0	17	0,29672	1	0
15	-1,91123	0	0	31	0,29672	1	1
6	-1,85098	0	0	14	0,31863	1	0
8	-1,73466	0	0	26	0,60277	1	0
11	-1,64871	0	0	54	0,71708	1	1
5	-1,49915	0	0	36	0,72041	1	1
24	-1,12013	0	0	33	0,72224	1	1
7	-1,07864	0	0	42	0,91245	1	1
13	-1,00020	0	0	40	0,94673	1	1
20	-0,93563	0	0	1	1,10260	1	0
27	-0,77018	0	0	32	1,12819	1	1
51	-0,73803	0	1	52	1,19384	1	1
21	-0,64960	0	0	37	1,25092	1	1
44	-0,63055	0	1	35	1,27546	1	1
2	-0,61966	0	0	50	1,27546	1	1
18	-0,57202	0	0	34	1,31137	1	1
28	-0,37283	0	0	48	1,34598	1	1
23	-0,24527	0	0	10	1,36146	1	0
9	-0,23802	0	0	47	1,44505	1	1
4	-0,22521	0	0	55	1,46466	1	1
3	-0,20785	0	0	45	1,47438	1	1
49	-0,20038	0	1	41	1,52136	1	1
29	-0,19984	0	1	46	1,56367	1	1
53	0,10937	1	1	56	1,62595	1	1
43	0,12364	1	1	39	2,79802	1	1

Fonte: Os autores.

Tabela 2 - Escores, classificação e situação de cada empresa

Classificação	N	%
Insolventes classificadas como solventes	6	21,4
Insolventes classificadas corretamente	22	78,6
Solventes classificadas como insolventes	4	14,3
Solventes classificadas corretamente	24	85,7

Fonte: Os autores.

Tabela 3 – Resultados da classificação do modelo

Para complementar optou-se por avaliar como ficariam os resultados ao se delimitar a região de penumbra. Os limites foram calculados de acordo com as Equações 3 e 4, sendo classificadas nesta zona as observações com escores $-0,6175 < Z < 0,4664$, correspondendo aos valores hachurados em cor cinza na Tabela 2. Nesta classificação a taxa de erro global cai para aproximadamente 8,9%. Os resultados estão dispostos na Tabela 4.

Classificação	N	%
Insolventes classificadas como solventes	2	3,6
Insolventes classificadas corretamente	16	28,6
Insolventes na região de “penumbra”	8	14,3
Solventes na região de “penumbra”	8	14,3
Solventes classificadas como insolventes	3	33,9
Solventes classificadas corretamente	19	5,3

Fonte: Os autores.

Tabela 4 – Resultados da classificação do modelo considerando a região de “penumbra”

Assim, esta proposta de modelo pode ser considerada uma opção para a continuidade dos estudos. Neste sentido seria recomendável a inclusão de mais amostras na análise e inclusão de uma amostra de validação. Ressalta-se aqui que o tamanho da amostra e o número e complexidade das variáveis usadas são fatores influentes para a definição do modelo (HAIR *et al*, 2005; CORRÊA; COSTA; MATIAS, 2006). Sendo assim, a inclusão proposta de mais elementos na amostra pode apresentar outro cenário resultante para a seleção das variáveis.

3. Conclusões e Considerações Finais

Este trabalho teve como objetivo desenvolver um modelo para previsão da insolvência de pequenas e micro empresas. Os resultados obtidos foram satisfatórios, considerando que este é um trabalho inicial, apresentando, no primeiro modelo desenvolvido, uma taxa de erro global de 17,86% para um modelo com 11 variáveis, na amostra de análise. Delimitou-se também uma região crítica (de penumbra), com duas linhas de corte, em que a solvência da empresa não pode ser prevista. Neste caso a taxa de erro global foi de 8,9%. O modelo proposto foi baseado na técnica de Análise Discriminante, a qual já foi amplamente utilizada em outros estudos também com o objetivo de analisar-se risco de crédito. A diferença é que, nos outros modelos publicados e até então conhecidos, os dados necessários para a análise são extraídos de relatórios emitidos somente para empresas de maior porte, enquanto que estes modelos utilizaram variáveis presentes no próprio sistema corporativo da instituição bancária e outras facilmente obtidas do próprio cliente, da praça em que atua e de sua atividade econômica. Reduz-se assim a burocracia no processo de crédito e no tempo de atendimento (ou não) dos seus pleitos. É importante lembrar-se que os modelos desenvolvidos foram realizados com base em informações cadastrais de micro e pequenas empresas de uma única região do país, o que limitou a quantidade de dados disponíveis para análise. Mesmo com esta limitação e através dos resultados obtidos, pôde-se verificar que é possível desenvolver-se modelos estatísticos eficientes para previsão de insolvência de micro e pequenas empresas. A partir dos resultados deste trabalho, espera-se que outros modelos mais eficientes possam ser desenvolvidos, através da redução da quantidade necessária de variáveis e da sinalização de outras que sejam ainda mais significativas para este tipo de análise. Outras técnicas estatísticas, como a regressão logística, e o aumento do tamanho amostral, também são opções para a continuidade dos trabalhos.

Referências

- ALTMAN, E. I.; SORRIS, T. A. financial early warning system for over-the-counter broker dealers. *Journal of Finance*, vol. 31, no.4, p. 1201 – 1277, 1976.
- ALTMAN, E. I. Predicting Performance in the Saving and Loan Association Industry. *Journal of Monetary Economics*. October 1977.
- ALTMAN, E. I, BAIDYA, T.K.N.; DIAS, L. M. R. Previsão de Problemas Financeiros em Empresas. *Revista de Administração de Empresas*, Vol 19, nº1, 1979.
- BNDES Banco Nacional de Desenvolvimento Social.** Porte da Empresa. Disponível em: http://www.bndes.gov.br/SiteBNDES/bndes/bndes_pt/Institucional/Apoio_Financeiro/porte.html. Acesso em: 10/09/2013.
- CORRÊA, A.C.C; COSTA, R.D.M.; MATIAS, A.B. Previsão de Insolvência de Pequenos Bancos Brasileiros. In: Seminários em Administração FEA-USP, 2006.
- ELISABETSKY, R. Um modelo matemático para decisões de crédito no banco comercial. 1976. 190 fls. Dissertação (Mestrado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 1976.
- HAIR JR, J.F.; ANDERSON, R.E.; TATHAN, R.L.; BLACK, W.C. *Análise Multivariada de Dados*. Porto Alegre: Bookman, 2005.

- MARIO, P.C.** Análise Discriminante. In: Corrar, L. J.; Paulo, E.; Dias Filho, J.M. (coordenadores) *Análise Multivariada para os cursos de administração, ciências contábeis e economia*. São Paulo: Editora Atlas, 2012.
- MATEUS, R.S.** Análise de insolvência empresarial: uma abordagem financeira fundamentalista com aplicação do método estatístico multivariado e da técnica discriminante. 2010. 79 fls. Dissertação (Mestrado) – Mestrado profissional em desenvolvimento regional e gestão de empreendimentos locais. Universidade Federal de Sergipe, 2010.
- MATIAS, A. B.; SIQUEIRA, J.O.** Risco Bancário: Modelo de Previsão de Insolvência de Bancos no Brasil. *Revista de Administração*, Vol. 31, n.2, 1996.
- PANDELÓ JR, D.R.; PANDELÓ, R.V.** Utilização da Análise Discriminante para previsão de insolvência bancária. *Revista Tecnológica de Crédito*, Vol. 66, 2008.
- PASOLD, B.A.** Previsão de Insolvência de Micro e Pequenas Empresas Utilizando a Técnica de Análise Discriminante. Trabalho de conclusão de curso. Graduação em Engenharia de Produção e Sistemas, Centro de Ciências Tecnológicas, Universidade do Estado de Santa Catarina, Joinville, 2010.
- PFAFF B.; MCNEIL, A.** 2012. QRM: Provides R-language code to examine Quantitative Risk. Management concepts. R package version 0.4-8. Disponível em: <http://CRAN.R-project.org/package=QRM>. Acesso em: 22/07/2013.
- PINHEIRO, L. E. T.; SANTOS, C.P.; COLAUTO, R.D.; PINHEIRO, J.L.** Validação de modelos brasileiros de previsão de insolvência. *Contabilidade Vista & Revista*, v.18, n. 4, p. 83-103, 2007.
- SANCHEZ, G.** 2013. DiscriMiner: Tools of the Trade for Discriminant Analysis. R package version 0.1-25. Disponível em: <http://CRAN.R-project.org/package=DiscriMine>. Acesso em: 22/07/2013.
- SECRIST, H.** National Bank Failures and non Failures. The Principia Press, 1938.
- SILVA, J.O.; WIENHAGE, P.; SOUZA, R.P.S.; LYRA, R.L.W.C.** Capacidade Preditiva de Modelos de Insolvência com base em Números Contábeis e Dados Descritivos. *REPEC Revista de Educação e Pesquisa em Contabilidade*, Vol. 6, n.3, p.246-261, 2012.
- SINKEY, J. F.** A Multivariate Statistical Analysis of the Characteristics of Problem Banks. *Journal of Finance*, v. 30, 1975.
- VENABLES, W. N.; RIPLEY, B. D.** *Modern Applied Statistics with S*. Fourth Edition. New York: Springer, 2002.

ANEXO A – Variáveis originais para a análise

Variável	Nomenclatura
Ramo de Atividade - Serviço	x1
Ramo de Atividade - Comércio	x2
Início de Operação (anos)	x3
Capital Social	x4
Faturamento Bruto Anual (Últ 12 meses)	x5
Tipo de Propriedade da empresa - próprio em nome dos sócios	x6
Tipo de Propriedade da empresa - na própria residência	x7
Tipo de Propriedade da empresa - alugado	x8
Pontualidade Fontes externas - não apurado	x9
Pontualidade Fontes externas - pontual	x10
Pontualidade pgto. junto a esta Instituição - pequenos atrasos (até 30 dias)	x11
Pontualidade pgto. junto a esta Instituição - últimos três anos ok	x12
Pontualidade pgto. junto a esta Instituição - sem experiência	x13
Cliente de outros bancos - sim	x14
Concentração de Vendas nos cinco principais clientes (%)	x15
Possui seguro empresarial - sim, outras seguradoras.	x16
Possui seguro empresarial - sim, na própria instituição	x17
Margem de lucro sobre as vendas	x18
Quantidade de cheques devolvidos nos últimos 12 meses	x19
Quantidade de cheques devolvidos nos últimos seis meses	x20
Qtde dias com excesso na conta corr. ou cheque especial nos ult. 12 meses	x21
Recursos Computáveis	x22
SFN - Endiv vencer acima de 360 dias ou a liberar	x23
SFN - Endiv a vencer em até 360 dias	x24
SFN - Endiv vencido há até 60 dias	x25
Tempo de conta (anos)	x26
Total de compras efetuadas nos últimos três meses	x27
Utilização média do ch especial ult seis meses	x28
Utilização média do ch especial ult 12 meses	x29
Valor das restrições SERASA ult cinco anos - não impeditivas	x30
Valor das restrições SERASA ult cinco anos - impeditivas	x31
Valor médio da folha de pagamento ult três meses	x32
Sócio - Conceito na praça - não apurado	x33
Sócio - Conceito na praça - bem conceituado	x34
Sócio - pontualidade fontes externas - não apurado	x35
Sócio - pontualidade fontes externas - bem conceituado	x36
Sócio - tempo de experiência no ramo de atividade (anos)	x37
Sócio - dedicação aos negócios - integral	x38
Sócio - quantidade de cheques devolvidos ult 12 meses	x39
Sócio - quantidade dias excesso na cc ou ch esp ult 12 meses	x40
Sócio - quantidade restrições ult. três anos	x41
Sócio - Recursos Computáveis	x42
Sócio - tempo de conta (anos)	x43
Sócio - Utilização média do ch esp ult. 12 meses	x44
Sócio - percentual de participação no capital da empresa	x45
Sócio 2 - Conceito na praça - não apurado	x46
Sócio 2 - Conceito na praça - bem conceituado	x47
Sócio 2 - pontualidade fontes externas - não apurado	x48
Sócio 2 - pontualidade fontes externas - bem conceituado	x49
Sócio 2 - tempo de experiência no ramo de atividade (anos)	x50
Sócio 2 - dedicação aos negócios – integral	x51
Sócio 2 - quantidade de cheques devolvidos ult 12 meses	x52
Sócio 2 - quantidade dias excesso na cc ou ch esp ult 12 meses	x53
Sócio 2 - Quantidade restrições ult 3 anos	x54
Sócio 2 - Recursos Computáveis	x55
Sócio 2 - tempo de conta (anos)	x56
Sócio 2 - Utilização média do ch esp ult 12 meses	x57
Sócio 2 - percentual de participação no capital	x58
Adimplência	y