



Classificação de Risco de Crédito: Um Comparativo entre Modelos de Análise Discriminante e Regressão Logística nas Empresas de Capital Aberto Brasileiras

Resumo: Os modelos de classificação de risco de crédito podem ser utilizados pelas empresas para prever uma situação financeira difícil com antecedência, evitando ou reduzindo os prejuízos financeiros. Também, contribuem para a reversão da insolvência, uma vez que um modelo de previsão eficaz pode prever problemas financeiros com um ano de antecedência, desta forma, a empresa possui tempo para tomada de decisões e ações necessárias. Sendo assim, o objetivo deste estudo é identificar o modelo mais adequado para classificação de risco de crédito das companhias abertas brasileiras, comparando modelos de análise discriminante e regressão logística. Foram utilizados como procedimentos metodológicos os métodos de pesquisa bibliográfica e documental com uma abordagem quantitativa. A amostra da pesquisa compreende 38 companhias abertas brasileiras. Os modelos escolhidos para validação foram os de Altman, Baidya e Dias; Matias; Brito e Assaf Neto e Scarpel e Milioni. Utilizou-se o *software SPSS Statistics* para realizar os testes de qualidade e significância dos modelos e a construção da curva ROC para validação. Os resultados do estudo indicam que a técnica de regressão logística apresenta um ótimo poder de discriminação, demonstrando uma área de 0,911 na curva ROC, podendo prever a insolvência com um ano de antecedência com um nível de 78,9% de precisão.

Palavras-chave: Risco de crédito; Insolvência; análise discriminante; regressão logística

Linha Temática: Gestão de Risco de Crédito



Contabilidade e Perspectivas Futuras

Florianópolis, SC, Brasil
Centro de Eventos da UFSC
12 a 14 de agosto de 2018

1. Introdução

O processo de gerenciamento de risco de crédito pode ser definido, de acordo com Cauoette, Altman e Narayanan (1998, tradução nossa), como a possibilidade de que a expectativa de recebimento de uma quantia, durante um determinado período, não se cumpra. Conforme Bessis (2011, tradução nossa), o risco de crédito pode ser entendido como as perdas geradas por um evento de *default* do tomador ou pela redução da sua qualidade de crédito. Sendo assim, há diversas situações que caracterizam um evento de *default*, como por exemplo, o atraso no pagamento de uma obrigação ou o início de um procedimento legal como a falência.

Diante disso, Minussi, Damacena e Ness Junior (2002) explicam que, ao aplicar um modelo de previsão de insolvência, estima-se que os problemas de uma companhia podem ser previamente identificados, uma vez que sua situação financeira tende a agravar, até atingir o estágio de falência.

Sendo assim, Martins (2003) afirma que os primeiros estudos relacionados à previsão de insolvência foram desenvolvidos por volta da década de 30, no entanto, somente a partir da década de 60, com o uso de técnicas estatísticas é que o tema ganhou impulso. Minussi *et al.* (2002) citam que existem diversas técnicas multivariadas para a construção de modelos de previsão de insolvência. Dentre elas, a análise discriminante e a regressão logística são as mais utilizadas.

Diante deste contexto, este estudo procura responder a seguinte questão de pesquisa: Qual o modelo mais adequado para classificação de risco de crédito das companhias abertas brasileiras utilizando um comparativo entre modelos de análise discriminante e regressão logística? Desta forma, obtém-se o objetivo desta pesquisa que consiste em verificar o modelo mais adequado para classificação de risco de crédito das companhias abertas brasileiras, comparando modelos de análise discriminante e regressão logística.

Para tanto, possui como objetivos específicos realizar o levantamento e diagnosticar a produção científica sobre risco de crédito, técnicas estatísticas e modelos de classificação de crédito, avaliar os modelos de classificação de risco de crédito a partir dos indicadores contábeis, aplicar as variáveis na amostra selecionada por meio dos testes de qualidade e significância e mensurar o poder discriminante dos modelos através da curva ROC, identificando o mais adequado.

De acordo com Stüpp (2015), a previsão de insolvência pode ser utilizada pelos usuários das empresas como uma ferramenta para analisar o risco de crédito e prever uma situação financeira difícil com antecedência, podendo evitar ou reduzir os prejuízos financeiros.

Um estudo realizado pela Serasa Experian identificou que no ano de 2016 houve um recorde histórico em recuperações judiciais requeridas com 1.770 solicitações, demonstrando um aumento de 44,8% em relação a 2015. O número de pedidos de falência também é preocupante, visto que neste ano cerca de 1.846 empresas entraram com o processo. (<http://noticias.serasaexperian.com.br/blog/2017/01/12/comercio-industria-e-servicos-batem-recorde-de-pedidos-de-recuperacao-judicial-em-2016-mostra-serasa-experian/>, recuperado em 04, agosto, 2017).

O setor de serviços registrou o maior número de pedidos de recuperação judicial do ano com 713 solicitações, demonstrando um aumento de 48,5% em relação a 2015. Em seguida, o setor de comércio apresentou 611 pedidos, tendo um crescimento de 51,2% em relação ao ano anterior, enquanto a indústria obteve um aumento de 24,2% com 446 pedidos de recuperação judicial. (<http://noticias.serasaexperian.com.br/blog/2017/01/12/comercio-industria-e-servicos-batem-recorde-de-pedidos-de-recuperacao-judicial-em-2016-mostra-serasa-experian/>, recuperado em 04, agosto, 2017).



Esta pesquisa justifica-se pela contribuição aos usuários das entidades, visto que visa garantir o modelo mais adequado para classificação de risco de crédito e, assim, auxiliar os gestores e investidores das empresas na tomada de decisões com uma ferramenta sólida. Também contribui para a reversão da insolvência, uma vez que, um modelo de previsão eficaz pode prever uma situação financeira difícil com um ano de antecedência.

Para tanto, a estrutura deste artigo está organizada, além desta introdução, em quatro seções: Na próxima seção, é apresentada a fundamentação teórica, que discorre sobre os principais pontos do risco de crédito, as técnicas estatísticas, os modelos de classificação de risco de crédito e estudos correlacionados. Já a terceira discorre sobre a metodologia utilizada no estudo: a pesquisa documental se deu por meio da análise das demonstrações contábeis das empresas, os testes de qualidade e significância foram utilizados para análise das funções e a curva ROC para validar os modelos analisados. Em seguida, a quarta seção apresenta a análise e validação dos modelos de previsão de insolvência. Finaliza-se com a conclusão da pesquisa e suas recomendações para estudos futuros.

2. Fundamentação Teórica

Neste tópico são apresentados os temas para o desenvolvimento deste estudo, tais como o risco de crédito, técnicas estatísticas, modelos de classificação de risco de crédito e estudos correlacionados, seus conceitos e suas aplicações.

2.1 Risco De Crédito

Segundo Brito e Assaf Neto (2008), o risco de crédito pode ser compreendido como a probabilidade de o credor ter prejuízos financeiros em razão de o tomador não quitar as obrigações assumidas nas condições acordadas. Sendo assim, Minussi *et al.* (2002) afirmam que a análise de risco de crédito requer a definição minuciosa de critérios que possam indicar a possibilidade de inadimplência do tomador.

De acordo com Davis (1999, tradução nossa), existem três modelos de análise de risco de crédito que buscam auxiliar os gestores na tomada de decisão de crédito, são eles: modelos preditivos, de risco e de *default*.

Os modelos preditivos são aplicados para previsão da capacidade de um cliente quitar suas obrigações, geralmente é utilizado para aprovação de um crédito ou para aumento de limite. Por outro lado, os modelos de risco são utilizados para prever se clientes que já fazem parte da carteira da empresa tendem a pagar ou estender uma possível inadimplência no pagamento, o principal objetivo deste modelo é analisar seu potencial de coleta. Por fim, os modelos de *default* são utilizados para prever se o cliente tem tendência à falência. (Davis, 1999, tradução nossa).

2.2 Técnicas Estatísticas

Conforme Mário (2002), busca-se verificar com a utilização de técnicas estatísticas se determinada empresa se classifica no grupo de solventes ou insolventes por meio da pontuação adquirida com base em dados obtidos das demonstrações contábeis. Assim, através do estudo realizado por Stüpp (2015), não há uma definição quanto à técnica estatística mais aplicada pelos autores para previsão de insolvência, contudo, dentre as mais utilizadas, destaca-se a análise discriminante e a regressão logística.

Nesse sentido, Brito e Assaf Neto (2008) afirmam que a regressão logística apresenta vantagens em relação à análise discriminante, uma vez que suas suposições iniciais são mais flexíveis. Pode-se observar que a análise discriminante está baseada em requisitos mais



limitados, como a normalidade das variáveis independentes e igualdade das matrizes de variância-covariância dos grupos.

2.2.1 Análise Discriminante

A análise discriminante é uma técnica estatística que procura definir dois ou mais grupos e classificar as observações em um deles com base em suas características individuais. É uma técnica utilizada principalmente para classificações nas quais a variável dependente é categórica, ou seja, qualitativa, como por exemplo: masculino ou feminino, solvente ou insolvente e as variáveis independentes são quantitativas. (Sanvincente & Minardi, 1998). Nessa perspectiva, Hair, Black, Babin, Anderson e Tatham (2009) ressalta que a análise discriminante busca realizar a combinação linear de duas ou mais variáveis independentes que discriminará da melhor forma os objetos nos grupos definidos inicialmente.

De acordo com Stüpp (2015), a variável estatística para a análise discriminante, também é denominada função discriminante, que é definida a partir de uma equação para ampliar a discriminação dos grupos. A equação é representada da seguinte forma: $Z = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$. Em que: Z é a variável dependente categórica, representada por uma pontuação; a é o intercepto da função; b é o coeficiente discriminante para a variável independente e X é o valor das variáveis independentes.

Barth (2002) explica que, após definida a função discriminante, deve-se definir o ponto de corte para utilizá-lo na classificação das observações iniciais. O ponto de corte é definido partindo do princípio de que, se Z for maior ou igual ao ponto de corte, a observação será classificada no grupo 1, por outro lado, se Z for menor que o ponto de corte, a observação será classificada no grupo 2.

2.2.2 Regressão Logística

A regressão logística, também conhecida como análise *logit*, é definida por Hair *et al.* (2009) como uma forma especializada de regressão formulada para prever e explicar uma variável categórica binária, e variáveis independentes métricas e não métricas.

De acordo com Barth (2002), o modelo de previsão utilizado na regressão logística é definido como: $\ln\left(\frac{p}{1-p}\right) = b_0 + b_1.X_1 + b_2.X_2 + b_3.X_3 + \dots + b_n.X_n$. Em que: p é a probabilidade de solvência; $(1 - p)$ é a probabilidade de insolvência; X são as variáveis preditoras; b são os coeficientes a serem atribuídos a cada uma das variáveis X .

Conforme Stüpp (2015), na regressão logística, os coeficientes estimam o efeito das alterações nas variáveis independentes sobre o logaritmo natural da razão de probabilidades, conhecido com *logit*. Para avaliar o impacto dos fatores sobre a probabilidade de solvência, eles devem ser transformados por meio de antilogaritmo, que é obtido pela expressão:

$$p = \frac{1}{1 + e^{b_0 + b_1.X_1 + b_2.X_2 + b_3.X_3 + \dots + b_n.X_n}}$$

Conforme Brito e Assaf Neto (2008), através dos resultados da regressão logística é possível medir a probabilidade de determinada empresa apresentar um estado de insolvência ou solvência, com base em um conjunto de características.

2.3 Modelos De Classificação De Risco De Crédito

Segundo Andrade (2004), existem três categorias de modelos de risco de crédito: modelos estocásticos de risco de crédito, modelos de risco de portfólio e modelos de



classificação de risco. Os modelos estocásticos buscam avaliar o comportamento estocástico do *default* e as variáveis que o determinam, os modelos de risco de portfólio têm por objetivo quantificar os riscos de uma carteira de crédito, medindo a distribuição de perdas ou de valor dela.

Por outro lado, os modelos de classificação de risco visam avaliar o risco de um devedor ou operação específica, atribuindo uma medida que corresponde à expectativa de ocorrência de *default*. Esse indicador pode ser expresso na forma de *rating* ou *scoring* (Andrade, 2004).

Chaia (2003) afirma que o objetivo do *scoring* é identificar os aspectos que definem a probabilidade de inadimplência de determinada empresa por meio da utilização de técnicas estatísticas. Por outro lado, Mendes (2014) afirma que a avaliação de *rating* contempla aspectos qualitativos, incluindo os fatores específicos da entidade e as características do setor, e os aspectos quantitativos que têm foco na análise financeira.

2.3.1 Modelo de Altman, Baidya e Dias

Altman, Baidya e Dias (1979) desenvolveram um modelo de previsão de insolvência utilizando a análise discriminante com uma amostra composta por 58 empresas, sendo 23 com problemas financeiros e 35 empresas de setores semelhantes e de porte equivalente às empresas da primeira amostra. O modelo obteve cerca de 84,2% de acerto na classificação das empresas e possui a seguinte equação: $Z = 1,44 + 4,03_{x2} + 2,25_{x3} + 0,14_{x4} + 0,42_{x5}$.

As variáveis são obtidas pelas seguintes fórmulas: $x2$: (não exigível - capital aportado pelos acionistas) / ativo total, $x3$: lucro antes dos juros e impostos / ativo total, $x4$: patrimônio líquido / exigível total, $x5$: vendas / ativo total. O ponto de corte desse modelo é 0. (Altman *et al.*, 1979).

2.3.2 Modelo de Matias

De acordo com Bez Birolo, Cittadin e Ritta (2011), o modelo de Matias foi desenvolvido em 1978 por meio da análise discriminante a partir de uma mostra de 50 empresas. O modelo é apresentado pela seguinte equação: $Z = 23,792_{x1} - 8,260_{x2} - 8,868_{x3} - 0,764_{x4} + 0,535_{x5} + 9,912_{x6}$. Em que: $x1$: patrimônio líquido/ativo total, $x2$: financiamentos e empréstimos bancários/ativo circulante, $x3$: fornecedores/ativo total, $x4$: ativo circulante/passivo circulante, $x5$: lucro operacional/lucro bruto, $x6$: disponível/ativo total.

2.3.3 Modelo de Brito e Assaf Neto

O modelo de previsão de insolvência de Brito e Assaf Neto (2008) foi desenvolvido com a utilização da regressão logística, tendo como amostra 30 empresas solventes e 30 insolventes. É apresentado pela seguinte equação: $\ln\left(\frac{p}{1-p}\right) = -4,740 - 4,528_{x12} + 18,433_{x16} - 414,080_{x19} - 11,028_{x22}$. Em que: $x12$: (lucros acumulados - reserva de lucros) /ativo total, $x16$: (passivo circulante financeiro + exigível a longo prazo financeiro) / ativo total, $x19$: (ativo circulante - passivo circulante) /ativo total, $x22$: (ativo circulante financeiro - passivo circulante financeiro) / vendas líquidas. O modelo obteve 88,3% de classificações corretas e possui ponto crítico 0.



Contabilidade e Perspectivas Futuras

Florianópolis, SC, Brasil
Centro de Eventos da UFSC
12 a 14 de agosto de 2018

2.3.4 Modelo de Scarpel e Milioni

O modelo de Scarpel e Milioni foi desenvolvido através da regressão logística a partir de uma amostra de 60 empresas brasileiras de capital aberto, a eficiência global foi de 87%. O modelo é apresentado pela seguinte equação: $\ln \left(\frac{p}{1-p} \right) = 0,190 + 3,312_{GAi} - 3,687_{EGi} + 10,997_{ROA AJIR}$. Em que: giro do ativo total (GA): receita/ativo total, endividamento geral (EG): exigível total/patrimônio líquido, retorno sobre o ativo total antes do pagamento de juros e imposto (ROA AJIR): lucro antes do pagamento de juros e imposto de renda/ativo total. (Scarpel & Milioni, 2002).

2.5 Estudos Correlacionados

Foram desenvolvidos muitos estudos para previsão de insolvência das empresas tendo como base indicadores extraídos das demonstrações contábeis. Mendes (2014) afirma que o primeiro estudo que se tem notícia sobre o tema é o de Fitzpatrick, desenvolvido em 1932. Esta pesquisa analisou 19 empresas solventes e insolventes, entre 1920 e 1929, comparando os índices dos dois grupos.

Taboada Pinheiro, Santos, Douglas Colauto e Lima Pinheiro (2007) cita que os estudos realizados sobre previsão de insolvência, tanto no Brasil como no exterior, fazem o uso de diferentes métodos estatísticos, sendo as principais técnicas utilizadas a Análise Discriminante Linear, Regressão Logística e Redes Neurais. O primeiro estudo sobre Modelos de Previsão de Insolvência utilizando o método da análise discriminante no Brasil foi proposto por Sthefen Charles Kanitz em 1974.

De acordo com Silva (1982) apresentou os resultados de um modelo desenvolvido em 1982, baseado na Análise Discriminante. Para operacionalizar o estudo a amostra foi composta por 419 empresas, divididas em empresas solventes e insolventes, sendo 337 empresas industriais e 82 comerciais. O modelo de Silva para a indústria obteve 90% de acerto na classificação de empresas solventes e 83% na classificação das insolventes. Já o modelo desenvolvido para empresas comerciais alcançou 90% de acerto na classificação das solventes e 72% na classificação das insolventes.

Em seguida, Sanvicente e Minardi (1998) utilizaram uma amostra constituída por 92 empresas com ações negociadas na Bolsa de Valores de São Paulo, dentre as quais 46 tiveram ações negociadas como concordatárias no período de 1986 a 1998. O modelo classificou corretamente 81,1% das concordatárias e 79,5% das não concordatárias.

Por fim, Minussi *et al.* (2002) construíram um modelo de previsão de insolvência com a utilização da regressão logística, a partir de uma amostra constituída por 323 empresas. A precisão do modelo foi de 94,85%.

3. Metodologia

Para desenvolvimento deste estudo, foram utilizados como procedimentos os métodos de pesquisa bibliográfica e documental. A pesquisa bibliográfica é realizada a partir do levantamento de referências teóricas publicadas por meios escritos e eletrônicos, como livros e artigos científicos. Este tipo de pesquisa permite conhecer o que já se estudou sobre o assunto. (Fonseca, 2002).

Por outro lado, a pesquisa documental, de acordo com Fachin (2006, p. 146) “Consiste na coleta, classificação, seleção e utilização de toda a espécie de informação. Compreende técnicas e métodos que facilitam a identificação”.



Assim, a pesquisa procede com uma abordagem quantitativa e sendo classificada quanto aos objetivos como um estudo exploratório. A pesquisa quantitativa, de acordo Gressler (2004), é caracterizada pelo desenvolvimento de hipóteses, definições de variáveis e quantificação nas modalidades de coleta de dados e informações, utilizando métodos estatísticos.

Gil (2010, p.27), afirma que “as pesquisas exploratórias têm como propósito proporcionar maior familiaridade com o problema, com vistas a torna-lo mais explícito ou a construir hipóteses. Seu planejamento tende a ser bastante flexível, pois interessa considerar os mais variados aspectos relativos ao fato ou fenômeno estudado”.

A população do estudo, a partir da qual a amostra foi selecionada, compreende as companhias abertas brasileiras que constam no cadastro da Comissão de Valores Mobiliários (CVM) em situação de recuperação judicial ou extrajudicial, contemplando 32 empresas.

Para seleção da amostra, o primeiro procedimento adotado foi a identificação das empresas que entraram com pedido de recuperação judicial entre os anos de 2011 a 2016, que possuem registro ativo na CVM e que não constava estado de falência no ano anterior ao pedido de recuperação judicial, ficando com 19 companhias no estado de recuperação judicial ou extrajudicial.

A partir da identificação das empresas insolventes, procedeu-se a seleção das empresas solventes que completaram a amostra. O grupo de solventes foi constituído selecionando-se para cada empresa insolvente, uma solvete de setor econômico e região igual.

Para aplicar os modelos de previsão de risco de crédito, foram analisadas as demonstrações financeiras de um ano antes do pedido de recuperação judicial das empresas selecionadas, abrangendo o período de 2010 a 2015. Desta forma, a amostra final utilizada no desenvolvimento deste estudo compreende 38 empresas, conforme Tabela 1.

Tabela 1. Empresas que compõem a amostra da pesquisa

| Empresas Insolventes | Empresas Solventes | Ano Evento | Ano Analisado |
|--|---|-------------------|----------------------|
| Cia Industrial Schlosser S.A. | Karsten S.A. | 2011 | 2010 |
| Buettner S.A. Ind. e Comercio. | Cia Hering. | 2011 | 2010 |
| Centrais Elétricas do Pará S.A – CELPA | Investco S.A. | 2012 | 2011 |
| Tecnosolo S.A. | CR2 Empreendimentos Imobiliários S.A. | 2012 | 2011 |
| Teka Tecelagem Kuehnrich S.A. | Dohler S.A. | 2012 | 2011 |
| Refinaria Pet Manguinhos S.A. | Petro Rio S.A. | 2013 | 2012 |
| GPC Participações S.A. | Elekeiroz S.A. | 2013 | 2012 |
| Mangels Industrial S.A. | Kepler Weber S.A. | 2013 | 2012 |
| Óleo e Gás Participações S.A. | Qgep Participações S.A. | 2013 | 2012 |
| OGX Petróleo e Gás S.A. | Petróleo Brasileiro S.A. – Petrobras. | 2014 | 2013 |
| Inepar Equipamentos e Montagens S.A. | Whirlpool S.A. | 2014 | 2013 |
| Inepar S.A Ind. e Construções. | Indústrias Romi S.A. | 2014 | 2013 |
| Fibam Cia Industrial. | Cia Siderurgica Nacional. | 2014 | 2013 |
| Eneva S.A. | Itapebi Geração de Energia S.A. | 2015 | 2014 |
| Const Sultepa S.A. | Unicasa Indústria de Móveis S.A. | 2015 | 2014 |
| Lupatech S.A. | Cobrasma S.A. | 2015 | 2014 |
| Wetzel S.A. | Tupy S.A. | 2016 | 2015 |
| Oi S.A. | Telefônica Brasil S.A. | 2016 | 2015 |
| Viver Incorporadora e Construtora S.A. | Cyrela Brazil Realty S.A. Emprs. e Parts. | 2016 | 2015 |

Fonte: Autores



De acordo com Barth (2002), para mensurar a qualidade do modelo de previsão é possível aplicá-lo à amostra de desenvolvimento. No entanto, ao testá-lo dessa maneira, pode-se concluir que seu desempenho é bom, contudo, em casos extremos ele pode funcionar bem somente sobre esta amostra. Sendo assim, para garantir a eficiência do modelo de previsão de insolvência é necessário testá-lo em uma amostra distinta da que foi utilizada para seu desenvolvimento.

Dessa forma, os modelos foram aplicados à amostra de validação com o auxílio do software *SPSS Statistics* em sua versão 20.0, sendo assim, foram realizados os testes de qualidade e significância dos modelos de Altman, Baidya e Dias; Matias; Brito e Assaf Neto; Scarpel e Milioni e construção da curva ROC para validação.

Os testes de qualidade dos modelos de análise discriminante foram realizados através do *eigenvalue* que mede a diferença entre os grupos, e a correlação canônica que mensura a associação entre as variáveis independentes e a dependente. O teste de significância foi efetuado por meio do *Lambda de Wilks*, e o nível de significância foi estabelecido de acordo com a escala de Fisher, que considera os valores abaixo de 0,05 como um nível muito forte. (Guimarães, 2006).

Dessa forma, para os modelos de regressão logística, os testes de qualidade e significância foram efetuados por meio do *Likelihood Value* que mensura a capacidade de prever a ocorrência de um evento, o R^2 de Cox & Snell e Nagelkerke que indicam o quanto que a variável dependente é explicada pelas variáveis independentes e o teste de Hosmer e Lemeshow que verifica as diferenças entre os grupos. (Ferreira, 2010).

Por fim, a técnica estatística utilizada para validação dos modelos foi a curva ROC (*Receiver Operating Characteristic*) que está baseada nos conceitos de especificidade e sensibilidade. De acordo com Barth (2002), a especificidade é a correlação entre o número de empresas insolventes corretamente classificados e o número total de empresas insolventes na amostra. Por outro lado, a sensibilidade é definida pela correlação entre o número de empresas solventes corretamente classificados e o número total de empresas solventes na amostra

4. Análise dos Dados

Após a seleção da amostra de empresas e o cálculo dos índices econômico-financeiros, procedeu-se ao cálculo dos modelos no software *SPSS Statistics*, no qual utilizou-se valor 0 para indicar o estado de solvência e 1 para o estado de insolvência.

Para analisar a qualidade da função discriminante, foram calculados os *eigenvalues* ou autovalores das funções estimadas. De acordo com Nóbrega (2010), o autovalor demonstra a diferença entre os grupos nas funções discriminantes, dessa forma, quanto mais afastados de 1, maiores serão as variações entre os grupos explicados pela função discriminante. Para esta análise, o resultado é de apenas uma função, correspondendo a 100% da variância explicada em termos de diferenças entre grupos.

O *SPSS Statistics* demonstra a correlação canônica junto com a análise do autovalor, assim, Maroco (2007) explica que a correlação canônica mede a relação entre o grupo de variáveis que identificam os grupos e as funções discriminantes. Nesse sentido, o percentual que explica a variação entre as variáveis dependentes e independentes é determinado elevando-se ao quadrado a correlação canônica, este valor também é conhecido como R^2 . (Protásio, Trugilho, Neves, & Vieira, 2012).

Tabela 2 - *Eigenvalues* – Altman, Baidya e Dias

| <i>Eigenvalue</i> | % Variância | Cumulativo % | Correlação Canônica | R^2 |
|--------------------|-------------|--------------|---------------------|-------|
| 0,171 ^a | 100 | 100 | 0,382 | 0,146 |

Fonte: Dados da pesquisa e Autores



Tabela 3 - *Eigenvalues* - Matias

| Eigenvalue | % Variância | Cumulativo % | Correlação Canônica | R ² |
|--------------------|-------------|--------------|---------------------|----------------|
| 0,450 ^a | 100 | 100 | 0,557 | 0,310 |

Fonte: Dados da pesquisa e Autores

Conforme a tabela 2, o modelo de Altman, Baidya e Dias apresentou um autovalor de 0,171, enquanto o modelo desenvolvido por Matias, demonstrado na tabela 3, foi de 0,450, dessa forma, pode-se perceber que o segundo modelo é o que mais se aproxima de 1, demonstrando a menor variação entre os grupos. Dessa forma, o modelo de Matias apresentou maior relação entre a função discriminante entre os grupos, apresentando uma correlação canônica de 0,557, elevando este valor ao quadrado, pode-se verificar que 31% da variável dependente é explicada pelo modelo, logo, o modelo de Altman, Baidya e Dias evidenciou uma baixa correlação canônica de 0,382 com R² de 14,6%.

A significância da função discriminante é analisada por meio do teste de *Lambda de Wilk's*, que demonstra o quanto a função separa as observações em grupos diferentes. O objetivo do teste é ignorar a hipótese nula, sendo p-valor igual a zero, sendo assim, foi utilizado um nível de confiança de 95%, e o nível de significância deve ser igual ou menor que 0,05. (Nóbrega, 2010).

Tabela 4 - *Wilk's Lambda* – Altman, Baidya e Dias

| <i>Wilk's Lambda</i> | Sig. |
|----------------------|-------|
| 0,854 | 0,251 |

Fonte: Dados da pesquisa e Autores

Tabela 5 - *Wilk's Lambda* – Matias

| <i>Wilk's Lambda</i> | Sig. |
|----------------------|-------|
| 0,690 | 0,056 |

Fonte: Dados da pesquisa e Autores

A tabela 4 demonstra os resultados do teste para o modelo de Altman, Baidya e Dias, no qual apresenta significância de 0,251 e 0,056 para a de Matias, conforme tabela 5, dessa forma, observa-se que nenhuma das duas funções é significativa na diferenciação entre os grupos.

Para avaliar a capacidade de previsão dos modelos de análise discriminante, construiu-se uma matriz de classificação para demonstrar a quantidade de empresas classificadas correta e incorretamente pelos modelos. Conforme demonstra a tabela 6, o nível de acerto do modelo desenvolvido por Altman, Baidya e Dias foi de 76,3%, tendo sido classificadas corretamente 29 empresas, por outro lado, de acordo com a tabela 7, o modelo de Matias classificou corretamente 81,6% das empresas, sendo 31 das 38 empresas da amostra.

Tabela 6 - Matriz de classificação – Altman, Baidya e Dias

| Classificação | | Grupo Previsto | | Total |
|----------------|------------|----------------|------------|-------|
| | | Solvente | Insolvente | |
| Grupo Original | Solvente | 16 | 3 | 19 |
| | Insolvente | 6 | 13 | 19 |
| | Solvente | 84,2% | 15,8% | 100% |
| | Insolvente | 31,6% | 68,4% | 100% |

a. 76,3% dos casos originais classificados corretamente

Fonte: Dados da pesquisa e Autores



Contabilidade e Perspectivas Futuras

Florianópolis, SC, Brasil
Centro de Eventos da UFSC
12 a 14 de agosto de 2018

Tabela 7 - Matriz de classificação – Matias

| Classificação | Grupo Previsto | | Total |
|----------------|----------------|------------|-------|
| | Solvente | Insolvente | |
| Grupo Original | Solvente | 16 | 3 |
| | Insolvente | 4 | 15 |
| | Solvente | 84,2% | 15,8% |
| | Insolvente | 21,1% | 78,9% |

a. 81,6% dos casos originais classificados corretamente

Fonte: Dados da pesquisa e Autores

A análise da qualidade dos modelos de regressão logística foi realizada por meio do *Likelihood Value*, dos pseudos R^2 e do teste de Hosmer e Lemeshow. O *Likelihood Value* é um indicador da qualidade do ajustamento aos dados, quanto menor o seu valor, melhor é o ajustamento. Para medir o poder de explicação do modelo, são utilizados os Pseudos R^2 de Cox e Snell e Nagelkerke, que demonstram os efeitos que as variáveis independentes causam na variável dependente. (Maroco, 2007; Zanini, 2007)

Para este estudo será utilizado o pseudo R^2 de Nagelkerke, que se assemelha ao coeficiente de determinação da regressão linear, sendo que quanto maior for esta medida, melhor o ajuste do modelo. Cabe ressaltar, que o pseudo R^2 de Nagelkerke tem escala de zero a um, por outro lado, o pseudo R^2 de Cox e Snell é limitado e não alcança o valor máximo de um. (Silva Brito, Assaf Neto, & Corrar, 2009).

Tabela 8 - Resumo do modelo – Brito e Assaf Neto

| -2 Log Likelihood | Cox & Snell R^2 | Nagelkerke R^2 |
|---------------------|-------------------|------------------|
| 27,749 ^a | 0,481 | 0,641 |

Fonte: Dados da pesquisa e Autores

Tabela 9 - Resumo do modelo – Scarpel e Milioni

| -2 Log Likelihood | Cox & Snell R^2 | Nagelkerke R^2 |
|---------------------|-------------------|------------------|
| 48,035 ^a | 0,115 | 0,153 |

Fonte: Dados da pesquisa e Autores

Observa-se na tabela 8, que o -2 Log de *Likelihood* do modelo de Brito e Assaf Neto é 27, 749, enquanto o modelo de Scarpel e Milioni é 48,035, conforme demonstra a tabela 9. Assim, pode-se observar que as variáveis independentes do modelo de Brito e Assaf Neto explicam melhor a variável dependente, uma vez que é o mais próximo de zero. Da mesma forma, pode-se ainda dizer que o modelo de Brito e Assaf Neto possui um melhor ajustamento, visto que o R^2 de Nagelkerke é maior do que o apresentado no modelo de Scarpel e Milioni.

Em seguida foi realizado o teste de Hosmer e Lemeshow, que estima a relação entre os valores reais e os previstos da variável dependente. Diante disso, uma menor a diferença entre a classificação observada e a prevista, indica um melhor ajuste do modelo, portanto o valor de qui-quadrado não deve ser significativo, ou seja, é necessário que ele seja maior do que 0,05. (Hair *et al.*, 2009).

Tabela 10 - Teste de Hosmer e Lemeshow – Brito e Assaf Neto

| Qui-quadrado | df | Sig. |
|--------------|----|-------|
| 5,158 | 8 | 0,741 |

Fonte: Dados da pesquisa e Autores



Contabilidade e Perspectivas Futuras

Florianópolis, SC, Brasil
Centro de Eventos da UFSC
12 a 14 de agosto de 2018

Tabela 11 - Teste de Hosmer e Lemeshow – Scarpel e Milioni

| Qui-quadrado | df | Sig. |
|--------------|----|-------|
| 12,926 | 8 | 0,114 |

Fonte: Dados da pesquisa e Autores

Conforme demonstra a tabela 10, a significância do modelo de Brito e Assaf Neto é 0,741, o que significa que o modelo é capaz de produzir estimativas e classificações confiáveis. Em contrapartida, o modelo de Scarpel e Milioni apresenta uma significância de 0,114, de acordo com a tabela 11, estando acima do esperado, porém em um comparativo dos dois modelos, o primeiro apresenta resultados melhores do que o segundo.

Para análise dos modelos de regressão logística, também foram desenvolvidas as matrizes de classificação. Conforme demonstra a tabela 12, o modelo de Brito e Assaf Neto classificou corretamente 78,9% das empresas, totalizando 30 empresas. Já o nível de acerto do modelo de Scarpel e Milioni, de acordo com a tabela 13, foi de 73,7%, sendo 28 das 38 empresas da amostra classificadas corretamente.

Tabela 12 - Matriz de classificação – Brito e Assaf Neto

| Classificação | | Grupo Previsto | | Total |
|----------------|------------|----------------|------------|-------|
| | | Solvente | Insolvente | |
| Grupo Original | Solvente | 15 | 4 | 19 |
| | Insolvente | 4 | 15 | 19 |
| | Solvente | 78,9% | 21,1% | 100% |
| | Insolvente | 21,1% | 78,9% | 100% |

a. 78,9% dos casos originais classificados corretamente

Fonte: Dados da pesquisa e Autores

Tabela 13 - Matriz de classificação – Scarpel e Milioni

| Classificação | | Grupo Previsto | | Total |
|----------------|------------|----------------|------------|-------|
| | | Solvente | Insolvente | |
| Grupo Original | Solvente | 16 | 3 | 19 |
| | Insolvente | 7 | 12 | 19 |
| | Solvente | 84,2% | 15,8% | 100% |
| | Insolvente | 36,8% | 63,2% | 100% |

a. 73,7% dos casos originais classificados corretamente

Fonte: Dados da pesquisa e Autores

Por fim, para medir a capacidade de discriminação de todos os modelos, foi desenvolvido a curva ROC. Hosmer e Lemeshow (2000, apud CABRAL, 2013), apresentam uma avaliação geral do resultado da área sob a curva ROC: área do intervalo entre 0,7 e 0,8 apresentam uma discriminação aceitável, já uma área entre 0,8 e 0,9 significa uma excelente discriminação, por fim, a área acima de 0,9 representa uma excepcional discriminação.



Contabilidade e Perspectivas Futuras

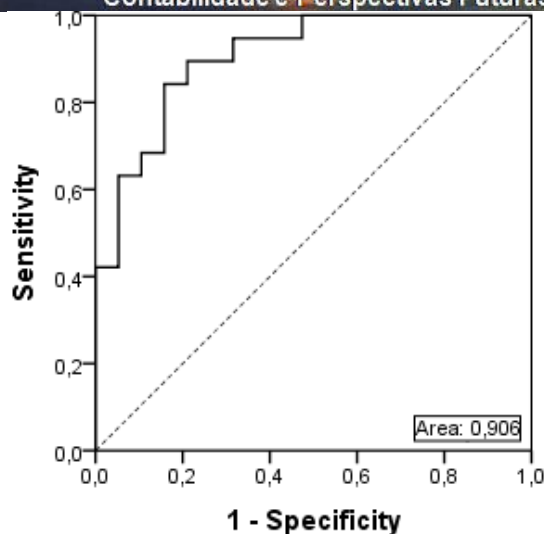


Figura 1. Curva ROC – Altman, Baidya e Dias
 Fonte: Dados da pesquisa e Autores

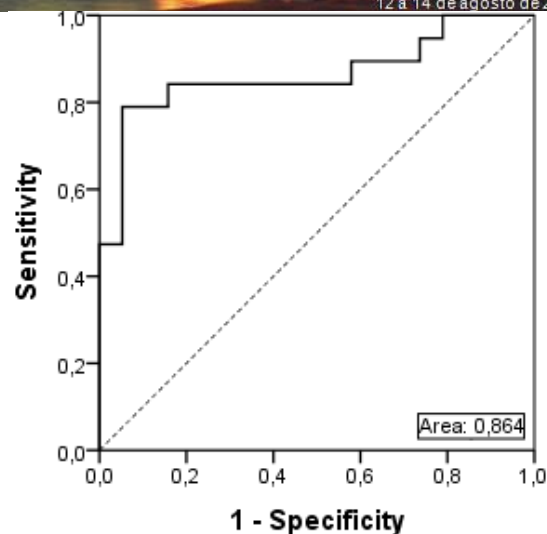


Figura 2. Curva ROC – Matias
 Fonte: Dados da pesquisa e Autores

Observa-se na figura 1, que o modelo de Altman, Baidya e Dias obteve uma discriminação excepcional, uma vez que a área sob a curva foi de 0,906, enquanto o modelo de Matias, demonstrado na figura 2, a área sob a curva foi de 0,864 representando uma discriminação excelente.

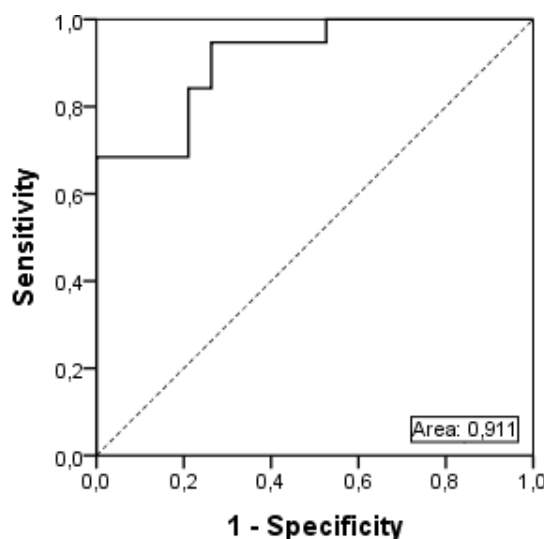


Figura 3. Curva ROC – Brito e Assaf Neto
 Fonte: Dados da pesquisa e Autores

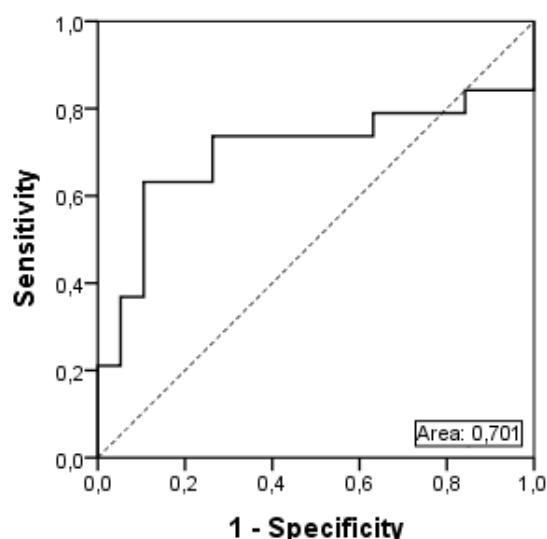


Figura 4. Curva ROC – Scarpel e Milioni
 Fonte: Dados da pesquisa e Autores

Por fim, a curva ROC do modelo desenvolvido por Brito e Assaf Neto, representada na figura 3, revela que a área sob a curva é de 0,911, indicando um excepcional poder de discriminação do modelo. Por outro lado, o modelo de Scarpel e Milioni na figura 4 demonstram uma discriminação aceitável, visto que a área sob a curva foi de 0,701.



5. Conclusão

O objetivo deste estudo foi verificar o modelo mais adequado para classificação de risco de crédito das companhias abertas brasileiras, comparando modelos de análise discriminante e regressão logística. Para tanto, utilizou-se dois modelos para cada técnica estatística escolhida.

Os modelos de risco de crédito desenvolvidos por meio da análise discriminante apresentaram um autovalor satisfatório, no entanto a correlação canônica e o R^2 ficaram abaixo do esperado, uma vez que 31% da variável dependente é explicada pelo modelo de Altman, Baidya e Dias, e apenas 15% é explicado pelo modelo de Matias.

No que se refere à significância das funções discriminantes, pode-se observar que nenhuma das funções é significativa na diferenciação entre grupos, visto que todas apresentaram valores maiores do que 0,05. De modo geral, os modelos demonstraram bons índices de classificação, dado que o modelo de Altman, Baidya e Dias obteve 76,3% de acerto, e o modelo de Matias atingiu 81,6% de classificações corretas.

Seguindo o processo, foi realizada a análise da qualidade dos modelos de regressão logística, desta forma, pode-se contar que o modelo desenvolvido por Brito e Assaf Neto obteve um melhor ajustamento, uma vez que apresentou um menor valor de $-2 \text{ Log de Likelihood}$ e maior pseudo R^2 . Os resultados do modelo de Scarpel e Milioni ficaram muito abaixo do esperado.

Em geral, os modelos de regressão logística apresentaram resultados satisfatórios para o teste de Hosmer e Lemeshow, no entanto, o modelo de Brito e Assaf Neto demonstrou maior capacidade de produzir estimativas e classificações confiáveis. Pode-se confirmar essa observação através da matriz de classificação, na qual o modelo de Brito e Assaf Neto obteve 78,9% de acerto, e o modelo de Scarpel e Milioni classificou corretamente 73,7% das empresas.

Sendo assim, conclui-se que a técnica de regressão logística apresenta um ótimo poder de discriminação, visto que o modelo de Brito e Assaf Neto apresentou resultados excelentes em todas as análises, além disso, esta técnica possui uma maior flexibilidade, uma vez que é possível utilizar variáveis de natureza métrica e não métrica.

Verifica-se que o modelo de Scarpel e Milioni apresentou resultados abaixo do esperado pois foram utilizadas apenas três variáveis independentes, que por sua vez não possuíam um bom nível de significância e qualidade. No estudo realizado pelos autores, o modelo desenvolvido alcançou 87% de classificações corretas, no entanto quando aplicado em outra amostra, não obteve o mesmo nível.

Por fim, como recomendações para pesquisas futuras, propõe-se realizar o comparativo com outras técnicas multivariadas como as redes neurais, também aplicar uma quantidade maior de modelos de previsão de insolvência e utilizar o método de Lachembruch para validação dos modelos.

Referências

Altman, E. I., Baidya, T. K., & Dias, L. M. R. (1979). Previsão de problemas financeiros em empresas. *Revista de administração de empresas*, 19(1), 17-28.

Altman, E. I., & Hotchkiss, E. (2010). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (Vol. 289). John Wiley & Sons.



Andrade, F. W. M. D. (2004). Desenvolvimento de modelo de risco de portfólio para carteiras de crédito a pessoas físicas (Doctoral dissertation).

Barth, N. L. (2002). Métodos de discriminação entre grupos: aplicação ao problema da concessão de crédito (Doctoral dissertation).

Bessis, J. (2015). Risk management in banking. John Wiley & Sons.

Bez Birolo, P., Cittadin, A., & de Oliveira Ritta, C. (2011). Análise de crédito por meio de modelos de previsão de insolvência: um estudo de caso na Empresa Cerâmica Alfa SA. Revista Catarinense da Ciência Contábil, 10(29).

Silva Brito, G. A., & Assaf Neto, A. (2008). Modelo de classificação de risco de crédito de empresas. Revista Contabilidade & Finanças-USP, 19(46).

Silva Brito, G. A., Assaf Neto, A., & Corrar, L. J. (2009). Sistema de classificação de risco de crédito: uma aplicação a companhias abertas no Brasil. Revista Contabilidade & Finanças-USP, 20(51).

Cabral, C. I. S. (2013). Aplicação do modelo de regressão logística num estudo de mercado (Doctoral dissertation).

Caouette, J. B., Altman, E. I., & Narayanan, P. (1998). Managing credit risk: the next great financial challenge (Vol. 2). John Wiley & Sons.

Chaia, A. J. (2003). Modelos de gestão do risco de crédito e sua aplicabilidade ao mercado brasileiro (Doctoral dissertation, Universidade de São Paulo).

Davis, K. T. (1999). Credit Scoring: A Brief Look. BUSINESS CREDIT-NEW YORK-, 101, 32-34.

Fachin, O. (2006). Fundamentos de metodologia. In Fundamentos de metodologia.

Ferreira, F. S. (2010). Aplicação da técnica de regressão logística na previsão de insolvência no segmento das cooperativas de crédito no Brasil. Universidade Federal do Rio de Janeiro, Brazil (Doctoral dissertation, MSc Dissertation).

Fonseca, J. J. S. (2002). Metodologia da Pesquisa Científica.

Gil, A. C. (2010). Como elaborar projetos de pesquisa.

Gressler, L. A. (2004). Introdução à pesquisa. Edições Loyola.

Guimarães, A. (2006). Previsão de insolvência: um modelo baseado em índices contábeis com a utilização da análise discriminante. (Insolvency Predictor: A Model Based on Account Index with Application of Discriminant Analysis).

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). Análise multivariada de dados. Bookman Editora.



Mário, P. D. C. (2002). Contribuição ao estudo da solvência empresarial: uma análise de modelos de previsão-estudo exploratório aplicado em empresas mineiras (Doctoral dissertation, Universidade de São Paulo).

Maroco, J. (2007). Análise estatística com utilização do SPSS.

Martins, M. S. (2003). A previsão de insolvência pelo modelo de Cox: uma contribuição para a análise de companhias abertas brasileiras.

Matias, A. B. (1978). Contribuição às técnicas de análise financeira: um modelo de concessão de crédito. Trabalho apresentado ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. São Paulo.

Mendes, I. O. D. V. (2014). Variáveis discriminantes dos estágios de insolvência de Empresas.

Minussi, J. A., Damacena, C., & Ness Jr, W. L. (2002). Um modelo de previsão de solvência utilizando regressão logística. *Revista de Administração Contemporânea*, 6(3), 109-128..

NÓBREGA, D. (2010). Análise discriminante utilizando o software SPSS. 54 f. Monografia (Bacharelado em Estatística), Universidade Estadual da Paraíba, Campo Grande.

Taboada Pinheiro, L. E., Santos, C. P., Douglas Colauto, R., & Lima Pinheiro, J. (2007). Validação de modelos brasileiro de previsão de insolvência. *Contabilidade Vista & Revista*, 18(4).

Protásio, T. P., Trugilho, P. F., Neves, T. A., & Vieira, C. M. M. (2012). Análise de correlação canônica entre características da madeira e do carvão vegetal de *Eucalyptus*. *Scientia Forestalis*, 40(95), 317-326.

SANVICENTE, A. Z., & Minardi, A. M. A. F. (1998). Identificação de indicadores contábeis significativos para a previsão de concordata de empresas. Instituto Brasileiro de Mercado de Capitais, Working Paper.

Scarpel, R. A., & Milioni, A. Z. (2002). Utilização conjunta de modelagem econométrica e otimização em decisões de concessão de crédito. *Pesquisa Operacional*, 22(1), 61-72.

Silva, J. P. D. (1982). Modelos para classificação de empresas com vistas à concessão de crédito (Doctoral dissertation).

Stüpp, D. R. (2015). Previsão de insolvência a partir de indicadores contábeis: evidências de empresas listadas na BM&FBOVESPA nos anos 2004-2013.

Zanini, A. (2007). Regressão Logística e redes neurais artificiais: um problema de estrutura de preferência do consumidor e classificação de perfis de consumo. Faculdade de Economia Aplicada FEA/UFJF: Universidade Federal de Juiz de Fora.