

# MODELO DE CLASSIFICAÇÃO DE RISCO DE CRÉDITO DE GRANDES EMPRESAS

## **Autores**

**GIOVANI ANTONIO SILVA BRITO**

Universidade de São Paulo

**ALEXANDRE ASSAF NETO**

FEA - Ribeirão Preto

## **Resumo:**

O mercado financeiro tem presenciado nos últimos anos um movimento de evolução nas técnicas de avaliação e gerenciamento de risco de crédito. Nesse contexto, destaca-se uma maior ênfase por parte das instituições financeiras na utilização de modelos quantitativos de risco de crédito. No presente estudo, é desenvolvido um modelo de classificação de risco de crédito de grandes empresas e testado o conteúdo informacional das demonstrações contábeis em relação a eventos de insolvência empresarial. Utilizando a técnica estatística da regressão logística, o modelo de risco de crédito foi construído com base em uma amostra de empresas de capital aberto classificadas como solventes e insolventes no período entre 1994 e 2004. As variáveis explicativas do modelo são índices calculados a partir das demonstrações contábeis, utilizados para caracterizar a situação econômico-financeira da empresa. Para validar os resultados do modelo, foi utilizado o método das múltiplas sub-amostras e uma Curva ROC. Os resultados alcançados mostram que o modelo proposto possibilita a previsão da ocorrência de eventos de default com significativo índice de acerto e evidenciam que as demonstrações contábeis fornecem informações que permitem a classificação das empresas como prováveis solventes ou prováveis insolventes.

## **1. Introdução**

O processo de avaliação e gerenciamento de risco de crédito em instituições financeiras vem passando por um movimento de evolução ao longo dos últimos anos. Os métodos tradicionais de decisão baseados em critérios julgamentais têm perdido espaço nas atividades de crédito dos bancos, que buscam instrumentos mais eficazes para mensurar o risco dos tomadores e das carteiras de crédito.

Nesse contexto, observa-se uma maior ênfase das instituições na utilização de modelos quantitativos como suporte às decisões de concessão de crédito e à gestão das carteiras. Os modelos de risco de crédito compõem um ferramental técnico que supre de informações os gestores, contribuindo para que estes tomem decisões que atendam às diretrizes estabelecidas nas políticas de crédito da instituição.

O objetivo desta pesquisa é desenvolver um modelo de classificação de risco de crédito de grandes empresas que atuam no Brasil, utilizando a técnica estatística da regressão logística. O escopo do modelo é prever a ocorrência de um evento de default com a empresa no horizonte de tempo de um ano, visando principalmente a subsidiar os gestores no processo de concessão de crédito.

O modelo proposto estabelece uma relação estatística entre o default da empresa e um conjunto de índices econômico-financeiros calculados a partir das demonstrações contábeis. Com base nessa relação, o estudo avalia se as demonstrações contábeis fornecem informações que permitam aos seus diversos usuários prever a ocorrência de uma insolvência empresarial.

O restante deste artigo está organizado da seguinte maneira. Na seção 2 são abordados conceitos sobre risco de crédito. Na seção 3 são descritos os principais tipos de modelos de risco de crédito. A seção 4 inclui a parte empírica da pesquisa, onde é apresentado o modelo de classificação de risco desenvolvido. Por fim, a seção 5 exibe as conclusões e considerações finais do estudo.

## **2. Risco de Crédito**

O conceito de crédito pode ser analisado sob diversas perspectivas. Para uma instituição financeira, crédito se refere principalmente à atividade de colocar um valor à disposição de um tomador de recursos sob a forma de um empréstimo ou financiamento, mediante compromisso de pagamento em uma data futura.

O crédito geralmente envolve a expectativa do recebimento de um valor em um certo período de tempo. Nesse sentido, Caouette *et al.* (1999, p.1) afirmam que o risco de crédito é a chance de que esta expectativa não se cumpra. De forma mais específica, o risco de crédito é entendido como a possibilidade de o credor incorrer em perdas, em razão de as obrigações assumidas pelo tomador não serem liquidadas nas condições pactuadas.

Segundo Bessis (1998, p.81), o risco de crédito pode ser definido pelas perdas geradas por um evento de default do tomador ou pela deterioração da sua qualidade de crédito. Há diversas situações que podem caracterizar um evento de default de um tomador. O autor cita como exemplo o atraso no pagamento de uma obrigação, o descumprimento de uma cláusula contratual restritiva (*covenant*), o início de um procedimento legal como a concordata e a falência ou, ainda, a inadimplência de natureza econômica, que ocorre quando o valor econômico dos ativos da empresa se reduz a um nível inferior ao das suas dívidas, indicando que os fluxos de caixa esperados não serão suficientes para liquidar as obrigações assumidas.

A deterioração da qualidade de crédito do tomador não resulta em uma perda imediata para a instituição financeira, mas sim no incremento da probabilidade de que um evento de default venha a ocorrer. Nos sistemas de classificação de risco, as alterações na qualidade de crédito dos tomadores dão origem às chamadas migrações de risco. Destaca-se que cada instituição financeira adota seu próprio conceito de evento de default, estando normalmente relacionado ao atraso no pagamento de um compromisso assumido pelo tomador por períodos como 60 ou 90 dias.

O risco de crédito pode ser avaliado a partir dos seus componentes, os quais compreendem o risco de default, o risco de exposição e o risco de recuperação. O risco de default está associado à probabilidade de ocorrer um evento de default com o tomador em um certo período de tempo, o risco de exposição decorre da incerteza em relação ao valor do crédito no momento do default, enquanto que o risco de recuperação se refere à incerteza quanto ao valor que pode ser recuperado pelo credor no caso de um default do tomador.

O risco de recuperação depende do tipo do default ocorrido e das características da operação de crédito, como valor, prazo e garantias. O risco de default é também tratado por “risco cliente”, pois está vinculado às características intrínsecas do tomador de crédito. Os riscos de exposição e de recuperação são tratados por “risco operação”, uma vez que estão associados a fatores específicos da operação de crédito.

A mensuração de risco de crédito é o processo de quantificar a possibilidade de a instituição financeira incorrer em perdas, caso os fluxos de caixa esperados com as operações de crédito não se confirmem. O risco de default constitui a principal variável deste processo, podendo ser definido como a incerteza em relação à capacidade do devedor honrar os seus compromissos assumidos.

### 3. Modelos de Risco de Crédito

Os modelos de risco de crédito compreendem ferramentas e aplicações que têm por objetivo principal mensurar o risco de tomadores e transações individuais ou de uma carteira de crédito como um todo. Segundo Andrade (2005, p.1), os modelos de risco de crédito podem ser classificados em três grupos: modelos de classificação de risco, modelos estocásticos de risco de crédito e modelos de risco de portfólio.

Os modelos de classificação de risco buscam avaliar o risco de um tomador ou operação em específico, atribuindo uma medida que representa a expectativa de risco de default, geralmente expressa na forma de uma classificação de risco (*rating*) ou pontuação (*score*). Os modelos de classificação de risco são utilizados pelas instituições financeiras em seus processos de concessão de crédito.

Os modelos estocásticos de risco de crédito são aqueles que têm por objetivo avaliar o comportamento estocástico do risco de crédito ou das variáveis que o determinam. Esses modelos são utilizados pelas instituições financeiras principalmente para precificar<sup>1</sup> títulos e derivativos de crédito.

Os modelos de risco de portfólio visam a estimar a distribuição estatística das perdas ou do valor de uma carteira de crédito, a partir da qual são extraídas medidas que quantificam o risco daquele portfólio. Esses modelos constituem uma importante ferramenta no processo de gestão de riscos das instituições, pois permitem que o risco de crédito seja avaliado de forma agregada, considerando os efeitos da diversificação produzidos pelas correlações entre os ativos da carteira. Os modelos de risco de portfólio também são utilizados para cálculo do capital econômico a ser alocado pela instituição.

Entre os modelos de classificação de risco, tem sido objeto de especial atenção por parte de pesquisadores os chamados modelos de previsão de insolvência. Os modelos de previsão de insolvência são aqueles que têm por objetivo principal medir a probabilidade de uma empresa incorrer em um evento de default ao longo de um certo período de tempo. Esses modelos são construídos a partir de uma amostra de casos históricos de empresas tomadoras de crédito, dividida em dois grupos: um que engloba as que incorreram em eventos de default, chamadas de insolventes, e outro que compreende as que não incorreram em default, chamadas de solventes.

A partir das características das empresas da amostra, são identificadas as variáveis que melhor discriminam as empresas que se tornaram insolventes e as que permaneceram solventes no período analisado. O conjunto de variáveis selecionadas é então utilizado para classificar as empresas proponentes de novas operações crédito como prováveis solventes ou prováveis insolventes.

Um tipo de variável tradicionalmente utilizada para discriminar empresas solventes e insolventes são índices econômico-financeiras calculados a partir das demonstrações contábeis. Índices são relações entre contas ou grupos de contas das demonstrações contábeis, que têm por objetivo evidenciar determinados aspectos da situação econômica e financeira da empresa.

A utilização de índices econômico-financeiros como variáveis explicativas em modelos de risco de crédito fundamenta-se no conceito de que o evento de default geralmente não é um processo abrupto e imprevisível. A deterioração da situação econômico-financeira

---

<sup>1</sup> Nesta pesquisa, emprega-se o neologismo “precificar”, derivado do termo inglês “*pricing*”, com o sentido de “apreçar”, uma vez que aquele termo é amplamente utilizado no mercado financeiro.

da empresa tende a ocorrer de forma gradual, levando, em última instância, à degradação completa da sua qualidade de crédito e ao default. Como os índices evidenciam a deterioração da situação econômico-financeira da empresa ao longo do tempo, eles podem ser utilizados para prever a ocorrência do default.

Os modelos de previsão de insolvência geralmente se baseiam em técnicas estatísticas de análise multivariada, como regressão linear, análise discriminante e regressão logística. Mais recentemente, outras técnicas têm sido utilizadas no desenvolvimento de modelos de risco de crédito, como redes neurais e algoritmos genéticos.

Os modelos que buscam prever a insolvência de empresas têm sido estudados há várias décadas nos meios acadêmicos. Algumas pesquisas se tornaram referência na literatura, como por exemplo os trabalhos de Beaver (1966), Altman (1968) e Ohlson (1980) em âmbito internacional, e os de Kanitz (1976), Matias (1978) e Altman *et al.* (1979) no Brasil.

## **4. Estudo Empírico**

### **4.1. Amostra de Empresas**

O objetivo central deste estudo é desenvolver um modelo de classificação de risco de crédito de grandes empresas. Na construção de um modelo de risco de crédito, o primeiro elemento a ser definido é o tipo de evento de default que se deseja prever. O conceito de default adotado nesta pesquisa é o início de um procedimento legal de concordata ou falência. Assume-se que quando uma empresa se torna concordatária ou falida, a instituição credora incorre em perdas com a operação de crédito. Esse conceito de default foi utilizado em razão de ele atender aos propósitos do estudo, bem como de se tratar de uma informação disponível publicamente.

O modelo de risco de crédito foi desenvolvido a partir de uma amostra composta por empresas categorizadas como solventes e insolventes. O termo insolvente é utilizado para fazer referência a uma empresa que iniciou um processo de concordata ou falência durante o período analisado. Para os objetivos desta pesquisa, não há diferença entre a concordata e a falência, ambas caracterizando uma situação de insolvência.

A população de estudo a partir da qual a amostra foi selecionada compreende as empresas de capital aberto não financeiras, com papéis listados na Bolsa de Valores do Estado de São Paulo (Bovespa). Para seleção da amostra, o primeiro procedimento adotado foi a identificação das empresas insolventes. A partir dos relatórios Boletim Diário de Informações – BDI e Suplemento de Orientação, publicados pela Bovespa, foram identificadas as empresas que tiveram ações negociadas como concordatárias no período compreendido entre os anos de 1994 e 2004. Além das concordatárias, também foram classificadas como insolventes as empresas que constaram do cadastro de companhias abertas da Comissão de Valores Mobiliários (CVM) na situação de falidas durante o mesmo período.

Definiu-se como ano de ocorrência do evento de default, ou simplesmente ano do evento, aquele em que a empresa passou a ter suas ações negociadas como concordatária ou que passou a constar do cadastro da CVM como falida.

A partir da identificação das empresas insolventes, procedeu-se à seleção das empresas solventes que completaram a amostra. O grupo de solventes foi constituído selecionando-se, para cada empresa insolvente incluída na amostra, uma empresa solvente do mesmo setor econômico e de tamanho equivalente, segundo o valor dos seus ativos. A classificação setorial adotada foi a utilizada pela Economática.

A amostra final utilizada no desenvolvimento do modelo compreendeu 60 empresas, sendo 30 insolventes que se tornaram concordatárias ou falidas entre 1994 e 2004, e 30 solventes que foram emparelhadas com as primeiras. Esse método de emparelhamento da amostra foi baseado nos trabalhos de Beaver (1966), Altman (1968) e Sanvicente e Minardi (1998), podendo ser considerado como uma amostragem por julgamento (não probabilística).

Apesar de a quantidade de empresas da amostra não ser elevada, dois aspectos importantes devem ser citados. Em primeiro lugar, o número de defaults ocorridos com grandes empresas geralmente é reduzido. Entretanto, as perdas que podem ser incorridas pela instituição credora são significativas, uma vez que as operações com esse tipo de cliente geralmente têm valores elevados. Assim, o risco de crédito de grandes empresas é tradicionalmente considerado como de baixa probabilidade e de alto impacto.

Além desse aspecto, a população de estudo da qual a amostra foi selecionada engloba as empresas de capital aberto. Segundo dados obtidos junto à Bovespa, a quantidade média anual de companhias abertas entre 1994 e 2004 é inferior a 250. Assim, pode-se considerar que o tamanho da amostra é proporcional à população de estudo e é coerente com o tipo de tomador considerado da pesquisa. O quadro 1 relaciona as empresas que compõem a amostra, seus setores de atuação e o ano do evento de default.

**Quadro 1 - Amostra de Empresas**

<b>Nr.</b>	<b>Empresa Insolvente</b>	<b>Setor Econômico</b>	<b>Ano do Evento</b>	<b>Nr.</b>	<b>Empresa Solvente</b>
1	Sibra	Siderurgia	1994	31	Belgo Mineira
2	Ferro Ligas	Siderurgia	1994	32	Ferbasa
3	Propasa	Papel e Celulose	1995	33	Votorantim
4	Londrimalhas	Têxtil	1995	34	Karsten
5	Aquatec	Química	1995	35	Bombril
6	Mesbla	Comércio	1995	36	Lojas Renner
7	Montreal	Participações	1995	37	Trevisa
8	Inbrac	Eletroeletrônicos	1995	38	Trafo
9	Cetenco	Construção	1995	39	Sultepa
10	Cibran	Química	1996	40	Biobrás
11	Velonorte	Têxtil	1996	41	Schlosser
12	Iderol	Transporte	1996	42	Recrusul
13	Trufana	Têxtil	1996	43	Têxtil Renaux
14	Casa José Silva	Comércio	1997	44	Globex
15	Pará de Minas	Têxtil	1997	45	Pettenati
16	Corbetta	Curtume	1997	46	Vulcabrás
17	Glasslite	Brinquedos	1997	47	Estrela
18	SPSCS	Transporte	1997	48	Randon
19	Tectoy	Eletroeletrônicos	1998	49	Gradiente
20	Arapuã	Comércio	1998	50	Bompreço
21	Copas	Química	1999	51	Fertibrás
22	Aubos Trevo	Química	1999	52	Fosfertil
23	Itaunense	Siderurgia	1999	53	Gerdau
24	Lisamar	Alimentos	2000	54	Sadia
25	Lorenz	Alimentos	2000	55	Granóleo
26	Sharp	Eletroeletrônicos	2000	56	Electrolux
27	Sano	Construção	2002	57	Sondotécnica
28	Eucatex	Construção	2003	58	Duratex
29	Chapécó	Alimentos	2004	59	Avipal
30	Parmalat	Alimentos	2004	60	Leco

## 4.2. Variáveis Explicativas

As variáveis explicativas do modelo de classificação de risco de crédito compreendem índices utilizados para caracterizar a situação econômico-financeira das empresas, calculados a partir das suas demonstrações contábeis publicadas. O modelo proposto nesta pesquisa está fundamentado na relação entre a situação econômico-financeira da empresa e a ocorrência do evento de default. Dessa forma, quanto pior for a situação econômico-financeira de uma empresa, maior será a probabilidade de ela incorrer em um evento de default.

As demonstrações contábeis das empresas foram coletadas junto às bases de dados da Economática e da Comissão de Valores Mobiliários. Quando disponíveis, foram utilizadas as demonstrações consolidadas. Para as empresas insolventes, foram obtidos os três últimos demonstrativos contábeis publicados anteriormente ao ano do evento e, para as empresas solventes, foram coletados os demonstrativos referentes aos mesmos exercícios utilizados nas empresas insolventes com as quais foram emparelhadas.

Há um grande número de índices econômico-financeiros que potencialmente prevêm a ocorrência de um default. Em razão de inexistir uma teoria consolidada sobre este assunto, foram testados 25 índices freqüentemente utilizados em estudos dessa natureza. Não foi objeto do trabalho a identificação de novos indicadores com poder de previsão de insolvência.

Os quadros 2 e 3 apresentam, respectivamente, os 25 índices econômico-financeiros e a notação utilizada em suas fórmulas de cálculo.

**Quadro 2 - Índices Econômico-Financeiros**

<b>Código</b>	<b>Índice</b>	<b>Fórmula</b>
X1	Liquidez geral	$(AC + RLP) / (PC + ELP)$
X2	Liquidez corrente	$AC / PC$
X3	Liquidez seca	$(AC - ESTOQUES) / PC$
X4	Liquidez imediata	$DISPONÍVEL / PC$
X5	Retorno sobre o patrimônio líquido	$LL / PL \text{ inicial}$
X6	Retorno sobre o ativo	$LAJIR / AT$
X7	Retorno sobre vendas	$LL / VL$
X8	Giro do ativo	$VL / AT$
X9	Margem operacional	$LAJIR / VL$
X10	Lucro operacional sobre despesas financeiras	$LAJIR / DF$
X11	Patrimônio líquido sobre ativo	$PL / AT$
X12	Lucros retidos sobre ativo	$(PL - CS) / AT$
X13	Patrimônio líquido sobre exigível total	$PL / (PC + ELP)$
X14	Endividamento total	$(PC + ELP) / AT$
X15	Endividamento de curto prazo	$PC / AT$
X16	Endividamento financeiro	$(PCF + ELPF) / AT$
X17	Imobilização do patrimônio líquido	$AP / PL$
X18	Estoques sobre ativo	$ESTOQUES / AT$
X19	Capital de giro líquido	$(AC - PC) / AT$
X20	Necessidade de capital de giro	$(ACO - PCO) / AT$
X21	Saldo de tesouraria sobre ativo	$(ACF - PCF) / AT$
X22	Saldo de tesouraria sobre vendas	$(ACF - PCF) / VL$
X23	Fluxo de caixa operacional sobre ativo	$FCO / AT$
X24	Fluxo de caixa operacional sobre exigível total	$FCO / (PC + ELP)$
X25	Fluxo de caixa operacional sobre endiv. financeiro	$FCO / (PCF + ELPF)$

**Quadro 3 - Notação das Fórmulas de Cálculo dos Índices**

Notação			
AC	Ativo circulante	FCO	Fluxo de caixa das operações
ACF	Ativo circulante financeiro	LAJIR	Lucro antes dos juros e imposto de renda
ACO	Ativo circulante operacional	LL	Lucro líquido
AP	Ativo permanente	PC	Passivo circulante
AT	Ativo total	PCF	Passivo circulante financeiro
CS	Capital social	PCO	Passivo circulante operacional
DF	Despesas financeiras	PL	Patrimônio líquido
ELP	Exigível a longo prazo	RLP	Realizável a longo prazo
ELPF	Exigível a longo prazo financeiro	VL	Vendas líquidas

Os índices econômico-financeiros utilizados no estudo compreendem indicadores de liquidez, de rentabilidade, de atividade, de estrutura, de análise dinâmica e de fluxo de caixa. A seguir, apresenta-se uma breve descrição desses indicadores. Informações detalhadas sobre índices econômico-financeiros podem ser obtidas em Assaf Neto (2002) e Silva (2003).

Os índices X1, X2, X3 e X4 são indicadores de liquidez que medem capacidade de pagamento, comparando direitos realizáveis e exigibilidades. Os índices X5, X6, X7 e X9 são indicadores de rentabilidade que avaliam os resultados gerados pela empresa. O índice X8 é um indicador de atividade que expressa a relação entre as vendas e os investimentos. O índice X10 é um indicador de cobertura de juros que mede a capacidade da empresa pagar esses encargos. O índice X11 mede a parcela dos ativos que é financiada por recursos próprios. O índice X12 indica o volume de lucros retidos pela empresa em relação ao ativo. O índice X13 mede a proporção de recursos próprios em relação aos recursos de terceiros na estrutura de capital.

Os índices X14, X15 e X16 são indicadores de estrutura que avaliam o grau de endividamento da empresa. O índice X17 mede a parcela dos recursos próprios que está comprometida com o ativo permanente. O índice X18 indica o volume de estoques da empresa em relação ao seu ativo total. Os índices X19, X20, X21 e X22 são indicadores de análise dinâmica que avaliam a situação financeira da empresa. Os índices X23, X24 e X25 são indicadores de fluxo de caixa que medem os recursos gerados pelas atividades operacionais.

Os índices foram calculados com base nos demonstrativos contábeis do penúltimo exercício anterior ao ano do evento de default. Não foram considerados os dados do último exercício, uma vez que a concordata de algumas empresas foi deferida antes que os demonstrativos tivessem sido publicados. Nesse caso, os índices do último exercício já poderiam estar refletindo a situação concordatária da empresa, o que prejudicaria a qualidade do modelo. Em relação a esse aspecto, Ohlson (1980, p.110) comenta que se o propósito do estudo é a previsão, utilizar os demonstrativos do último exercício não é um procedimento adequado. A utilização dos dados do penúltimo exercício no modelo assegura um prazo de antecedência de pelo menos um ano em relação ao evento de default.

As estatísticas descritivas dos índices econômico-financeiros calculados revelam que alguns indicadores apresentam médias entre as empresas solventes e insolventes bastante

<sup>2</sup> O índice X12 foi proposto originalmente por Altman (1968) e adaptado ao contexto brasileiro por Altman, Baidya e Dias (1977).

próximas. Para avaliar se as diferenças das médias são significantes estatisticamente, foi aplicado o teste de médias (teste-t). O teste de igualdade de médias para duas amostras tem por objetivo estimar se as médias de duas populações são iguais estatisticamente. Se as médias dos índices de solventes e insolventes forem iguais estatisticamente, o indicador não é relevante para modelo de risco de crédito.

O resultado do teste indica que, ao nível de significância de 5%, não pode ser rejeitada a hipótese nula de igualdade de médias para os índices X6, X8, X9, X18, X20, X23, X24 e X25. Os valores médios desses indicadores para os grupos de solventes e insolventes não apresentaram diferenças estatisticamente significantes, portanto eles não foram utilizados no modelo de risco de crédito. Os demais índices econômico-financeiros apresentaram diferenças de médias com significância estatística. A tabela 1 apresenta os resultados do teste:

**Tabela 1 – Teste de Igualdade de Médias para os Índices Econômico-Financeiros**

Índice	Teste-t de Igualdade de Médias	
	Estatística t	Valor p
X1	3,495	0,001
X2	5,549	0,000
X3	4,933	0,000
X4	4,769	0,000
X5	2,886	0,005
X6	0,534	0,595
X7	2,551	0,013
X8	-0,149	0,882
X9	0,253	0,801
X10	2,935	0,005
X11	3,900	0,000
X12	4,311	0,000
X13	4,211	0,000
X14	-3,662	0,001
X15	-4,168	0,000
X16	-6,642	0,000
X17	-3,424	0,001
X18	0,961	0,341
X19	5,673	0,000
X20	0,629	0,532
X21	8,347	0,000
X22	6,121	0,000
X23	1,416	0,162
X24	0,431	0,668
X25	0,274	0,785

#### 4.3. Técnica Estatística

O modelo de risco de crédito foi desenvolvido utilizando-se a técnica estatística da regressão logística. A regressão logística, ou análise logit, é uma técnica de análise multivariada, apropriada para as situações nas quais a variável dependente é categórica e assume um entre dois resultados possíveis (binária), tais como “normal ou anormal”, “cliente ou não cliente” e “solvente ou insolvente”.

O objetivo da regressão logística é gerar uma função matemática, cuja resposta permita estabelecer a probabilidade de uma observação pertencer a um grupo previamente determinado, em razão do comportamento de um conjunto de variáveis independentes. Os

coeficientes estimados pelo modelo de regressão indicam a importância de cada variável independente para a ocorrência do evento.

A regressão logística apresenta certas vantagens em relação à análise discriminante linear, principalmente devido às suas suposições iniciais serem menos rígidas. A análise discriminante linear está baseada em uma série de pressupostos bastante restritivos, como a normalidade das variáveis independentes e a igualdade das matrizes de variância-covariância dos grupos de interesse. Essas suposições podem não ser válidas em muitas situações práticas de análise de risco de crédito, principalmente quando há variáveis independentes de natureza não métrica. Segundo Hair *et al.* (1998, p.276), a regressão logística não assume esses rígidos pressupostos e é uma técnica bem mais robusta quando eles não são atendidos.

Outro aspecto que favorece a utilização da regressão logística é que seus resultados podem ser interpretados em termos de probabilidade. Esse fator se mostra particularmente importante nos modelos de risco de crédito, pois possibilita que seja medida a probabilidade de um determinado tomador assumir a condição de solvente ou insolvente, face um conjunto de atributos. Em relação à análise discriminante linear, Ohlson (1980, p.112) comenta que o resultado da equação discriminante é um escore que tem pouca interpretação intuitiva, pois ele é basicamente um dispositivo discriminatório de classificação ordinal.

Na regressão logística, há uma transformação na variável dependente, que é convertida em uma razão de probabilidades e posteriormente em uma variável de base logarítmica (transformação logística). Devido à natureza não linear dessa transformação, os coeficientes da regressão são estimados pelo método da máxima verossimilhança. O modelo da regressão logística assume a seguinte relação:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

onde “p” é a probabilidade de ocorrer o evento, “1-p” a probabilidade de não ocorrer o evento, “p/(1-p)” a razão de probabilidades, “Xi” as variáveis independentes e “bi” os coeficientes estimados.

Na regressão logística, os coeficientes medem o efeito de alterações nas variáveis independentes sobre o logaritmo natural da razão de probabilidades, chamado de logit. Para avaliar o impacto dos parâmetros sobre a probabilidade de ocorrer o evento, eles devem ser transformados por meio de antilogaritmo. A probabilidade associada à ocorrência do evento de interesse pode ser obtida pela expressão:

$$p = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k)}}$$

onde “e” é base dos logaritmos naturais (aproximadamente 2,718).

#### 4.4. Desenvolvimento do Modelo

Após a seleção da amostra de empresas, o cálculo dos índices econômico-financeiros e a definição da técnica de análise de dados, procedeu-se ao desenvolvimento do modelo. O objetivo deste procedimento é identificar uma relação entre as variáveis explicativas e o estado de solvência ou insolvência das empresas, de forma que o conjunto de índices possa ser utilizado para estimar a probabilidade de futuros eventos de default.

No modelo logit, a variável dependente pode assumir um valor entre zero e um. Neste estudo, foi atribuído o valor zero para indicar estado de solvência e o valor um para estado de insolvência. O ponto de corte do modelo é 0,5, portanto as empresas com resultado inferior a 0,5 são classificadas como solventes e as empresas com resultado superior a esse valor são classificadas como insolventes. As variáveis independentes testadas foram os 17 índices cujas médias entre solventes e insolventes apresentaram diferenças estatisticamente significantes.

O método utilizado para seleção das variáveis foi o *forward stepwise*, pelo critério do menor *Likelihood Ratio*. O método *stepwise* é baseado em um algoritmo estatístico que avalia a importância de cada variável independente e as inclui ou exclui do modelo segundo uma determinada regra. A importância de cada variável é definida em termos de uma medida de significância estatística do seu coeficiente. Os parâmetros utilizados foram 5% de significância para a entrada das variáveis e 10% para a saída. Também foram testados os métodos do maior coeficiente Wald e da maior probabilidade condicional de máxima verossimilhança, os quais produziram resultados idênticos.

O modelo final foi composto pelo intercepto e quatro variáveis explicativas. As variáveis incluídas no modelo foram os índices X12 (lucros retidos sobre ativo), X16 (endividamento financeiro), X19 (capital de giro líquido) e X22 (saldo de tesouraria sobre vendas). A função matemática do modelo é:

$$\ln\left(\frac{p}{1-p}\right) = -4,535 - 5,152.X_{12} + 19,069.X_{16} - 13,364.X_{19} - 17,488.X_{22}$$

Os coeficientes das variáveis apresentaram o sinal esperado. As variáveis X12, X19 e X22 possuem coeficientes negativos, indicando que quanto maiores os valores assumidos por esses índices, menor a probabilidade da empresa sofrer um default. Por outro lado, a variável X16 apresentou coeficiente positivo, mostrando que quanto maior o valor assumido por esse indicador, maior a probabilidade da empresa se tornar insolvente.

Um elemento importante a ser considerado nos modelos de regressão é a correlação entre as variáveis independentes. A inclusão de variáveis altamente correlacionadas não é desejável, pois estas variáveis, chamadas de colineares, fornecem informações similares para explicar o comportamento da variável dependente, prejudicando a capacidade preditiva do modelo. A tabela 2 apresenta os coeficientes de correlação de Pearson das variáveis do modelo. Conforme se verifica, os índices selecionados não são altamente correlacionados.

**Tabela 2 – Matriz de Correlação das Variáveis Explicativas**

	X12	X16	X19	X22
X12	1			
X16	-0,37	1		
X19	0,51	-0,45	1	
X22	0,54	-0,51	0,47	1

Para avaliar a capacidade de previsão do modelo, pode-se construir uma matriz de classificação, que mostra a quantidade de empresas classificadas correta e incorretamente pelo modelo. Conforme demonstra a matriz na tabela 3, o nível de acerto do modelo desenvolvido foi de 91,7%, tendo sido classificadas corretamente 55 das 60 empresas da amostra. Do grupo de solventes, 28 empresas foram classificadas corretamente e 2 classificadas erradamente, o que representa uma taxa de acerto de 93,3%. Em relação ao grupo de insolventes, houve 27 classificações corretas e 3 erradas, o que corresponde a uma taxa de acerto de 90%. O erro

tipo I, classificar uma empresa insolvente como solvente, foi de 10%, enquanto o erro tipo II, classificar uma empresa solvente como insolvente, foi de 6,7%.

**Tabela 3 – Matriz de Classificação do Modelo de Risco de Crédito**

Observado	Estimado		Total	Classificações Corretas
	Solventes	Insolventes		
Solventes	28	2	30	93,3%
Insolventes	3	27	30	90,0%
Total	31	29	60	91,7%

#### 4.5. Avaliação do Ajuste do Modelo

Para avaliar a qualidade de um modelo logístico, diversos testes e medidas podem ser utilizados. O primeiro passo é avaliar a significância das variáveis explicativas incluídas no modelo, o que é feito por meio de um teste Qui-Quadrado, sob a hipótese nula de que todos os coeficientes são iguais a zero. O resultado da aplicação deste teste no modelo desenvolvido teve estatística Qui-Quadrado de 59,771, com significância de 0,000, o que leva à rejeição da hipótese nula e à interpretação de que ao menos um dos coeficientes é estatisticamente diferente de zero.

A análise da significância estatística de cada coeficiente é feita com base no teste Wald. Semelhante ao teste t aplicado aos modelos lineares, o teste Wald avalia a hipótese nula de que o parâmetro estimado é igual a zero. A estatística Wald tem distribuição Qui-Quadrado, sendo calculada pelo quadrado da razão entre o coeficiente e o seu erro padrão. Os resultados do teste Wald para o modelo proposto indicam que os parâmetros das quatro variáveis independentes selecionadas são estatisticamente diferentes de zero. A tabela 4 apresenta as estimativas dos parâmetros das variáveis do modelo, bem como os erros padrão, as estatísticas Wald, os graus de liberdade e os valores de probabilidade do teste Wald.

**Tabela 4 – Coeficientes do Modelo de Risco de Crédito**

Variável	Coefficiente	Erro-Padrão	Teste Wald	g.l.	Valor-p
Intercepto	-4,535	2,154	4,431	1	0,035
X12	-5,152	2,793	3,403	1	0,065
X16	19,069	8,563	4,959	1	0,026
X19	-13,364	6,713	3,963	1	0,047
X22	-17,488	7,452	5,507	1	0,019

A avaliação do nível de ajuste de um modelo logístico pode ser realizada por meio do *Likelihood Value*, dos pseudo  $R^2$  e do teste de Hosmer e Lemeshow (1980). O *Likelihood Value* é uma medida da qualidade geral do modelo e quanto menor o seu valor, melhor é o ajuste do modelo como um todo. O Cox-Snell  $R^2$  e o Nagelkerke  $R^2$  são medidas que se assemelham ao coeficiente de determinação da regressão linear. Maiores valores dessas medidas indicam melhor ajuste do modelo. O Cox-Snell  $R^2$  baseia-se no *Likelihood Value* e tem uma escala que começa em zero mas não alcança um em seu limite superior. O Nagelkerke  $R^2$  é similar ao Cox-Snell  $R^2$ , porém tem uma escala vai de zero a um.

A tabela 5 apresenta os valores do *Likelihood Value*, Cox-Snell  $R^2$  e Nagelkerke  $R^2$  do modelo. Os dados evidenciam a redução do *Likelihood Value* e o aumento do Cox-Snell  $R^2$  e do Nagelkerke  $R^2$  proporcionado pela inclusão das variáveis explicativas em cada etapa do processo *stepwise*. As variáveis foram incluídas na seguinte seqüência: X16, X19, X22 e X12.

**Tabela 5 – Likelihood Value, Cox-Snell R<sup>2</sup> e Nagelkerke R<sup>2</sup>**

Step	Likelihood Value	Cox-Snell R <sup>2</sup>	Nagelkerke R <sup>2</sup>
1	50,103	0,424	0,565
2	36,454	0,541	0,721
3	27,761	0,603	0,804
4	23,407	0,631	0,841

O teste de Hosmer e Lemeshow (1980) avalia as diferenças entre as classificações previstas pelo modelo e as observadas. Se as diferenças forem significativas, o grau de acurácia do modelo não é bom. O teste divide os casos em classes e compara as frequências preditas e observadas em cada classe por meio de uma estatística Qui-Quadrado. A tabela 6 apresenta os resultados finais do teste. Ao nível de significância de 5%, aceita-se a hipótese nula de que não há diferenças significativas entre os valores preditos e observados, o que indica que o modelo é capaz de produzir estimativas e classificações confiáveis.

**Tabela 6 – Teste Hosmer e Lemeshow**

Step	Qui-Quadrado	g.l.	Valor-p
1	4,798	8	0,779
2	5,892	8	0,659
3	3,191	8	0,922
4	0,876	8	0,999

#### 4.6. Validação do Modelo

O modelo desenvolvido classificou corretamente 91,7% das empresas da amostra. No entanto, ao se testar o modelo com a própria amostra utilizada para a estimação dos seus parâmetros, pode-se concluir que o seu desempenho é bom quando, na realidade, ele pode funcionar bem apenas para estas observações. Assim, para avaliar se o modelo mantém o seu poder preditivo para outras amostras provindas da mesma população, são necessários testes para a sua validação. Segundo Hosmer e Lemeshow (2000, p.186), a validação do modelo é especialmente importante quando ele é usado com a finalidade de previsão de resultados.

A validação do modelo foi feita por meio do método baseado em múltiplas sub-amostras denominado *jackknife*, proposto por Lachenbruch (1967). O método *jackknife* é uma técnica estatística amplamente aceita na validação de modelos de classificação e é particularmente útil para pesquisas que envolvem amostras de tamanhos pequenos, pois permite que todas as observações sejam utilizadas na estimação dos parâmetros do modelo (NEOPHYTOU *et al.*, 2000, p.20).

O método está baseado no princípio “*leave-one-out*” e consiste em separar uma observação da amostra original, estimar os coeficientes do modelo com base no restante da amostra (n-1) e classificar a observação apartada utilizando a nova equação. O procedimento é repetido para toda a amostra (n vezes), de maneira que todas as observações sejam classificadas por modelos cujos parâmetros foram estimados com base nas demais. O percentual de classificações corretas é acumulado para todas as observações da amostra, indicando a precisão global do modelo. A tabela 7 apresenta a matriz de classificação obtida

com a aplicação do método *jackknife*. O percentual de acerto acumulado foi de 88,3%, tendo sido classificadas incorretamente sete empresas da amostra (11,7%), três do grupo de solventes e quatro do grupo de insolventes.

**Tabela 7 – Matriz de Classificação – Validação do Modelo**

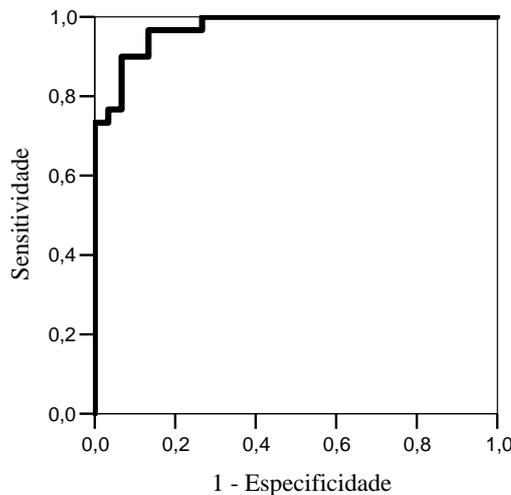
Observado	Estimado		Total	Classificações Corretas
	Solventes	Insolventes		
Solventes	27	03	30	90,0%
Insolventes	04	26	30	86,7%
Total	31	29	60	88,3%

Um segundo procedimento utilizado para avaliar a performance do modelo foi a construção de uma Curva ROC, conforme sugerido por Oliveira e Andrade (2002, p.3). A curva ROC (*Receiver Operating Characteristic*) constitui uma técnica bastante útil para validar modelos de risco de crédito e está baseada nos conceitos da sensibilidade e da especificidade. A sensibilidade é a proporção de acerto na previsão da ocorrência de um evento nos casos em que ele de fato ocorreu. A especificidade é proporção de acerto na previsão da não ocorrência de um evento nos casos em que ele de fato não ocorreu.

Para a construção da Curva ROC, são calculadas a sensibilidade e a especificidade para todas as observações da amostra, considerando diferentes pontos de corte do modelo. A curva é obtida registrando em um gráfico “sensibilidade” x “1 – especificidade” para os diversos pontos de corte. A área sob a curva mede a capacidade de discriminação do modelo. Hosmer e Lemeshow (2000, p.162) apresentam uma regra geral para avaliação do resultado da área sob a Curva ROC:

- a) área no intervalo entre 0,7 e 0,8: discriminação aceitável;
- b) área no intervalo entre 0,8 e 0,9: excelente discriminação;
- c) área acima de 0,9: excepcional discriminação.

A Curva ROC do modelo de risco de crédito, representada no gráfico 1, revela que a área sob a curva é de 0,972. Segundo a escala proposta pelos autores citados, esse valor indica um excepcional poder de discriminação do modelo.



**Gráfico 1 – Curva ROC**

## 5. Considerações Finais

O objetivo deste estudo foi desenvolver um modelo de classificação de risco de crédito de grandes empresas que atuam no Brasil. Utilizando a técnica estatística da regressão logística e um conjunto de quatro índices econômico-financeiros como variáveis explicativas, o modelo proposto possibilita a previsão da ocorrência de eventos de default com um ano de antecedência, alcançado significativo índice de acerto.

O modelo de risco de crédito desenvolvido estabelece uma relação entre o evento de default e a situação econômico-financeira da empresa, caracterizada por meio de índices calculados a partir das suas demonstrações contábeis. Com base nessa relação, foi possível testar o conteúdo informacional das demonstrações contábeis das empresas de capital aberto publicadas no Brasil. Concluiu-se que as demonstrações contábeis fornecem informações que permitem classificar as empresas como prováveis solventes ou prováveis insolventes, com bom nível de precisão.

Por fim, destaca-se que os modelos de risco de crédito derivados empiricamente, a exemplo dos modelos de classificação de risco, têm sido objeto de severas críticas nos meios acadêmicos, em razão de eles não estarem suportados por uma teoria explícita e bem desenvolvida. No entanto, embora esses modelos não tenham uma fundamentação teórica consistente, o expressivo nível de acerto nas classificações obtido de forma recorrente pelos trabalhos desenvolvidos nas últimas décadas sugere que a insolvência empresarial é um evento que pode ser previsto com sucesso. Em consequência, os modelos de risco de crédito desenvolvidos empiricamente ainda possuem um forte apelo prático, sobretudo pela sua aplicabilidade nas atividades de concessão e gerenciamento de risco de crédito das instituições financeiras.

## REFERÊNCIAS

- ALTMAN, Edward I. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. **The Journal of Finance**, v. 23, n. 4, p. 589-609, Sep. 1968.
- ALTMAN, Edward I. *et al.* Previsão de problemas financeiros em empresas. **Revista de Administração de Empresas**, v. 19, p. 17-28, Jan./Mar. 1979.
- ALTMAN, Edward I. *et al.* Zeta analysis: a new model to identify bankruptcy risk of corporations. **Journal of Banking and Finance**, v. 1, n. 1, p. 29-54, Jun. 1977.
- ALTMAN, Edward I.; SAUNDERS, Anthony. *Credit risk measurement: developments over the last 20 years*. **Journal of Banking & Finance**, v. 21, n. 11-12, p. 1721-1742, Dec. 1997.
- ANDRADE, Fabio W.M. Modelos de Risco de Crédito. **Tecnologia de Crédito**, n. 38, p. 23-53, 2003.
- ASSAF NETO, Alexandre. **Estrutura e análise de balanços**. 7. ed. São Paulo: Atlas, 2002.
- BASEL COMMITTEE ON BANKING SUPERVISION - BCBS. *Credit risk modelling: current practices and applications*. Basle: Bank for International Settlements, Apr. 1999.
- BEAVER, William H. *Financial ratios as predictors of failure*. **Journal of Accounting Research**, Empirical Research in Accounting: Selected Studies, v. 4, p. 77-111, 1966.
- BEAVER, William H. *Market prices, financial ratios, and the prediction of failure*. **Journal of Accounting Research**, v. 6, n. 2, p. 179-192, Autumn 1968.
- BESSIS, Joel. **Risk management in banking**. Chichester: John Wiley & Sons, 1998.
- BORGES, Olavo. Rating de crédito: considerações sobre os modelos. **Tecnologia de Crédito**, n. 24, p. 14-27, 2001.

- CAOQUETTE, John B. *et al.* **Gestão do risco de crédito: o próximo grande desafio financeiro.** Rio de Janeiro: Qualitymark, 1999.
- EISENBEIS, Robert A. *Pitfalls in the application of discriminant analysis in business, finance, and economics.* **The Journal of Finance**, v. 32, n. 3, p. 875-900, Jun. 1977.
- EISENBEIS, Robert A. *Problems in applying discriminant analysis in credit scoring models.* **Journal of Banking and Finance**, v. 2, n. 3, p. 205-219, Oct. 1978.
- HAIR Jr., Joseph F. *et al.* **Multivariate data analysis.** 5th ed. New Jersey: Prentice Hall, 1998.
- HOSMER, David W.; LEMESHOW, Stanley. *A goodness-of-fit test for the multiple logistic regression model.* **Communications in Statistics**, A10, p.1043-1069, 1980.
- HOSMER, David W.; LEMESHOW, Stanley. **Applied logistic regression.** 2nd ed. New York: John Wiley & Sons, 2000.
- KANITZ, Stephen C. **Indicadores contábeis e financeiros de previsão de insolvência: a experiência na pequena e média empresa brasileira.** São Paulo, 1976. Tese (Livre Docência) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo.
- LACHENBRUCH, P.A. *An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis.* **Biometrics**, v. 23, p. 639-645, Dec. 1967.
- MATIAS, Alberto B. **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, 1978.
- MERTON, Robert C. *On the pricing of corporate debt: the risk structure of interest rates.* **The Journal of Finance**, v. 29, n. 2, p. 449-470, May 1974.
- NEOPHYTOU, Evridiki *et al.* **Predicting corporate failure: empirical evidence for the UK.** University of Southampton, Working Paper, Sep. 2000.
- OHLSON, James A. *Financial ratios and the probabilistic predictions of bankruptcy.* **Journal of Accounting Research**, v. 18, n. 1, p. 109-131, Spring 1980.
- OLIVEIRA, José G.C.; ANDRADE, Fábio W.M. Comparação entre medidas de performance de modelos de *credit scoring*. **Tecnologia de Crédito**, n. 33, p. 35-47, 2002.
- PRESS, James S.; WILSON, Sandra. *Choosing between logistic regression and discriminant analysis.* **Journal of the American Statistical Association**, v. 73, n. 364, p. 699-705, Dec. 1978.
- SANVICENTE, Antonio Z.; BADER, Fani L. **Filing for financial reorganization in Brazil: event prediction with accounting and financial variables and the information content of the filing announcement.** Instituto Brasileiro de Mercado de Capitais, Working Paper, Mar. 1996.
- SANVICENTE, Antonio Z.; MINARDI, Andréa M.A.F. **Identificação de indicadores contábeis significativos para a previsão de concordata de empresas.** Instituto Brasileiro de Mercado de Capitais, Working Paper, Out. 1998.
- SAUNDERS, Anthony. **Medindo o risco de crédito.** Rio de Janeiro: Qualitymark, 2000b.
- SCOTT, James. *The probability of bankruptcy: a comparison of empirical predictions and theoretical models.* **Journal of Banking and Finance**, v. 5, n. 3, p. 317-344, Sep. 1981.
- SICSÚ, Abraham L. Desenvolvimento de um sistema de *credit scoring* – parte I. **Tecnologia de Crédito**, n. 4, p. 63-76, 1998a.
- SICSÚ, Abraham L. Desenvolvimento de um sistema de *credit scoring* – parte II. **Tecnologia de Crédito**, n. 5, p. 57-68, 1998b.

SILVA, José Pereira da. **Gestão e análise de risco de crédito**. 4. ed. São Paulo: Atlas, 2003.  
SPSS INC. *SPSS – Statistical Package for the Social Sciences*. Chicago, 2005.