

Anais do

VII Seminário Multidisciplinar ENIAC Pesquisa 2015

VII Encontro Da Engenharia Do Conhecimento Eniac

VII Encontro De Iniciação Científica Eniac

VII Fábrica de Artigos

MODELO ESTATÍSTICO PARA MENSURAR A PROBABILIDADE DE INADIMPLÊNCIA EM INSTITUIÇÕES FINANCEIRAS CONFORME O NOVO ACORDO DE BASILÉIA.

João Bosco Segreti

Ivam Ricardo Peleias

Marco Antonio Ferreira Lima

- Marco Antonio Ferreira Lima graduado em Estatística pela Ence (IBGE), especialista em Didática do Ensino Superior pelo Mackenzie e mestre em Contabilidade Estratégica pela Fecap. Possui vasta experiência no magistério em disciplinas de métodos quantitativos e finanças, tendo lecionando nos cursos de estatística, administração, contabilidade e tecnólogos, tanto na graduação como na pós-graduação Lato Sensu. Instituições nos quais exerci atividades: Uninove, Unip, Fecap, Unopec, Unifai, Faculdades São Luis. Leciona nas Faculdades Integradas Campos Salles. No ambiente não acadêmico, atua na área financeira e em pesquisa mercadológica, em empresas de médio a grande porte, nas áreas de marketing, produtos e modelagem de risco de crédito. Centro Universitário Álvares Penteado. UniFECAP. Mestrado em Controladoria e Contabilidade Estratégica. Avenida Liberdade, 532. Bairro Liberdade. CEP: 01502-001. São Paulo. SP Brasil. prof_marcolima@yahoo.com.br

- João Bosco Segreti. Centro Universitário Álvares Penteado. UniFECAP Mestrado em Controladoria e Contabilidade Estratégica. Avenida Liberdade, 532. Bairro Liberdade. CEP: 01502-001. São Paulo. SP Brasil.

- Ivam Ricardo Peleias. Centro Universitário Álvares Penteado. UniFECAP Mestrado em Controladoria e Contabilidade Estratégica. Avenida Liberdade, 532. Bairro Liberdade. CEP: 01502-001. São Paulo. SP Brasil.

RESUMO

Modelo estatístico para mensurar a probabilidade de inadimplência em instituições

financeiras conforme o novo acordo de Basiléia.

Este artigo apresenta um modelo para mensurar a probabilidade de inadimplência em crédito constituído por empresas junto a instituições financeiras. É um exercício aplicativo para o cálculo da probabilidade de inadimplência, uma componente da perda esperada, presente no modelo denominado Método Interno Próprio, em discussão no Novo Acordo da Basiléia. A pesquisa foi realizada com base em uma amostra de 500 empresas de vários ramos de atividades, do segmento varejo de um banco com os resultados efetivamente ocorridos. Conclui-se que o modelo proposto atende o objetivo nacional. O modelo, desenvolvido com a aplicação da Regressão Logística Múltipla, e uso do SPSS, é função de indicadores obtidos das demonstrações financeiras da amostra. Os resultados obtidos foram satisfatórios ao se comparar os percentuais de inadimplência apurados na amostra com a probabilidade do cliente cumprir ou não com suas obrigações creditícias.

Palavras-chaves: Probabilidade de inadimplência, demonstrações financeiras, regressão logística múltipla, risco de crédito, Novo Acordo da Basiléia.

ABSTRACT

Statistic model to measure the probability of default in Financial Institutions according to the New Basel Capital Accord. This article presents a model to measure the probability of default (PD) in credit constituted by companies joint to financial institutions. It constitutes an applicative exercise to calculate the PD, a component of the expected loss, which is present in the Internal Ratings Based Approach model, under discussion in the New Basel Capital Accord. The research was conducted on a sample of 500 companies of various sectors of activities, which are part of the retail segment of a Brazilian bank. The model, developed with the application of Multiple Logistic Regression and with the use of SPSS, is a function of indicators obtained from the financial statements. The results are gratifying, if compared to the percentage of default hissed in the sample with the occurred findings. As a conclusion, the model proposed reach the goal of measuring the probability of the clients accomplishing or not their obligations regarding credit.

Keywords: Probability of default, financial statements, multiple logistic regression, credit risk, New Basel Capital Accord.

INTRODUÇÃO

O Comitê da Basiléia sobre Fiscalização Bancária, consciente da necessidade de melhor gestão de risco dos bancos, propôs uma nova forma de gerar suficiência no nível de capital das instituições financeiras (IF's), para promover segurança e

efetividade do sistema. A proposta em discussão, quando concluída, substituirá o Acordo de 1988^[1]. O Novo Acordo de Capital da Basileia (*New Based Capital Accord - NBCA*) fundamenta-se em três pilares não mutuamente exclusivos:

Pilar 1 – Necessidades mínimas de capital:

- *Risco de crédito*: dispõe que as instituições bancárias podem escolher modos de alocar capital, seja pelo uso da abordagem padronizada, seja pela abordagem baseada em classificações internas – IRB (*Internal Ratings – Based Approach*). Instituições bancárias que adotarem o método padronizado ficarão dependentes de classificações de risco fornecidas por agências de *rating*^[2], enquanto as que optarem pelo IRB terão a prerrogativa de calcular, pelo menos, a probabilidade de inadimplência (*Probability to default - PD*).
- *Risco operacional*: também faculta as instituições bancárias a escolha de um entre três métodos: o básico, o padronizado e o mais sofisticado, denominado *Advanced Measurement Approach* (AMA).

Pilar 2 – Processo de exame da fiscalização:

Neste, foram estabelecidos quatro princípios que os supervisores deverão seguir na revisão de bancos, incluindo-se a avaliação da qualidade de gestão e a adequação dos planos de capital, com a possibilidade de intervenção se os níveis de capital do banco caírem abaixo do nível exigido.

Pilar 3 – Disciplina de mercado:

Este pilar define as exigências de divulgação para o mercado a respeito dos riscos do banco.

Dentre os pontos abrangidos pelo Novo Acordo de Capital da Basileia, o enfoque deste trabalho é no risco de crédito, na medida em que se propõe a apresentar um modelo para mensurar a inadimplência de crédito. O objetivo da pesquisa foi verificar se os dados contábeis fornecidos por empresas solicitantes de crédito a uma instituição bancária, convertidos em índices, são variáveis suficientemente robustas para serem empregadas em modelos quantitativos destinados à mensuração do risco na concessão de crédito. Trata-se assim, de calcular a probabilidade de inadimplência (*Probability of Default - PD*), condicionada à situação na qual as variáveis explicativas do modelo são os índices financeiros.

O problema de pesquisa, a hipótese da pesquisa, as análises efetuadas por analistas de crédito, sem o uso de modelos estatísticos ou econométricos, embutem em seu resultado aspectos de julgamento do profissional que a elaborou. Este fator pode transformar-se em aumento dos riscos de crédito, pois muitas vezes o julgamento realizado deixa de captar aspectos importantes na avaliação do crédito. Pelo fato de serem análises individuais, em muitas circunstâncias acabam não indicando o comportamento (*behavior*) do cliente com tendência à inadimplência.

Para mitigar este risco, as normas emanadas pelo Acordo da Basileia conduzem a modelos estatísticos. Neste cenário, os índices obtidos com a realização da análise das demonstrações contábeis (mais conhecida

como análise de balanços), podem se constituir na base de um modelo quantitativo.

Para testar se um modelo estatístico é mais eficiente do que o modelo baseado no julgamento do analista de crédito, a hipótese a ser verificada é a de que os métodos quantitativos proporcionam uma padronização no processo de avaliação do crédito, e geram uma resposta não intuitiva à análise, pois a mensuração do risco de crédito passa a ser resultado da aplicação de um modelo quantitativo. Esse modelo, fundamentado em métodos estatísticos, é independente, portanto, do grau de experiência de um analista de crédito, e apto a servir de ferramental para a obtenção das provisões de crédito.

O referencial teórico neste item será abordado os aspectos teóricos envolvendo o Novo Acordo da Basileia, e do modelo estatístico escolhido para o cálculo da probabilidade de inadimplência.

1 – COMPONENTES DE CRÉDITO NO NOVO ACORDO DA BASILÉIA.

O Novo Acordo de Capital da Basileia estabelece que as instituições financeiras dispõem, como métodos para alocação de capital em operações de crédito, da abordagem padronizada, e da fundamentada em classificações internas (IRB). A abordagem padronizada, mais simples de ser aplicada, é em parte sucessora das regras do Acordo de 1988, com algumas características adicionais. A abordagem padronizada estabelece ponderações fixas a cada uma das categorias

de risco, crédito, mercado e operacional, e usa informações de classificação de risco de agências de *rating*. Assim, a instituição financeira que adotar tal método não terá responsabilidade ou controle sob suas informações.

Na abordagem fundamentada em classificações internas (*Internal Ratings – Based Approach – IRB*) a instituição financeira usa dados disponíveis em seus sistemas de informação, para obter elementos que indiquem o nível de risco de seus clientes. Esta abordagem é composta de quatro grandes componentes de risco de crédito:

Probability to default -PD: probabilidade de inadimplemento do tomador;

Exposure at default - EAD: exposição a inadimplemento de uma operação de crédito;

Loss given default - LGD: Perda devido a inadimplemento de uma operação de crédito;

Effective maturiy - M: Tempo de vencimento.

As instituições financeiras que adotarem o método IRB terão que escolher entre o método fundamental e o avançado. A diferença entre um e outro é quanto à franquia que o NBCA fornece ao cálculo das componentes do risco de crédito. A instituição financeira que empregar o método fundamental somente poderá calcular a PD internamente. As demais componentes serão fornecidas pelos órgãos reguladores. O quadro 1 resume as fontes de informações que geram as diferenças metodológicas entre os modelos IRB:

Quadro 1– fontes de informações segundo o Modelo IRB

Modelo IRB		
Componentes	Fundamental	Avançado
PD	Instituição Financeira	Instituição Financeira
LGD	Estimativas padronizadas do órgão regulador	Instituição Financeira
EAD	Estimativas padronizadas do órgão regulador	Instituição Financeira
M	Estimativas padronizadas do órgão regulador	Instituição Financeira

Fonte: adaptação dos autores, baseada nos relatórios “Visão geral do novo Acordo de Capital da Basileia” e “O método baseado em Classificações Internas”. Comitê de Supervisão Bancária da Basileia (2001).

Do exposto, verifica-se que a abordagem padronizada é a que apresenta maiores limitações às instituições financeiras. As previsões quanto ao uso de métodos indicam que, internacionalmente, pelos menos no início, os bancos adotarão a forma padronizada. Contudo no Brasil, em função do perfil do mercado bancário, concentrado em pelo menos cinco bancos privados e dois federais, o Banco Central do Brasil (BACEN) incentiva o desenvolvimento do modelo IRB avançado.

Das componentes de risco apresentadas, a PD é única utilizada no modelo IRB –fundamental, e também a mais importante. Grandes modelos comerciais criados, tais como *KMV*, *CreditMetrics* e *CreditRisk+*, não analisados neste trabalho, foram desenvolvidos para a sua mensuração.

O modelo proposto neste trabalho é empregado com frequência nas situações denominadas *Credit Scoring*³. Estes modelos usam informações teóricas ou estatísticas para isolar os efeitos das diferentes características dos solicitantes de crédito quando ocorrem situações de atraso e inadimplência. Sua aplicação envolve o emprego de técnicas de classificação, que discriminam entidades de crédito em diferentes grupos de riscos. Modelos com esta característica permitem à instituição financeira:

- Ranquear as solicitações de crédito em termos de risco;
- Obter informações quanto a provável *performance* de uma particular solicitação de crédito;
- Obter estimativas das probabilidades de *defaults* necessárias aos modelos agregativos de risco de crédito das IF's;
- Desenvolver critérios objetivos de aprovação e rejeição de crédito.

A importância de modelos desta natureza pode ser medida pelo exposto constante do relatório emitido pelo Bank of International Settlements (2003, p.140), denominado *Consultative document of the New Basel Capital Accord*, que define as condições para o Pilar 2:

- Alocar capital eficientemente para risco de crédito;
- Desenvolver metodologias internas para avaliação de risco individual e agregado;
- Criar sistemas de *rating* interno como ferramenta de monitoração do risco de crédito.

Em termos de aplicação no mercado bancário nacional, o cálculo da PD, com o auxílio de modelos de *score*, fornece condições de classificação de risco de crédito de tal sorte que atendam aos dispositivos regulatórios da Resolução CMN nº 2.682/99. Esta determina que a operação de crédito seja classificada por níveis de risco, e orienta a elaboração da provisão de crédito em função

deste nível. O Quadro 2 apresenta, em resumo, as classificações e o valor mínimo de

provisionamento definidos na Resolução CMN nº 2.682/99.

Quadro 2 – dias em atraso e valor a provisionar, segundo o nível de risco

Nível de risco	Dias em atraso	Porcentagem do valor da operação a provisionar
AA	-	-
A	-	0,5 %
B	15 e 30	1,0 %
C	31 e 60	3,0 %
D	61 e 90	10,0 %
E	91 e 120	30,0 %
F	121 e 150	50,0 %
G	151 e 180	70,0 %
H	Maior que 180	100,0 %

Fonte: adaptação dos autores, baseada na Resolução CMN nº 2.682/99.

Quanto ao método de mensuração de risco, o BACEN, deixa a escolha cargo das instituições financeiras. Desta forma, não há impedimentos legais para a mensuração de risco de crédito (PD) com o uso de modelos de *Credit Scoring*.

2. O MODELO DE CÁLCULO DA PROBABILIDADE DE INADIMPLÊNCIA

O método empregado para o cálculo da probabilidade de inadimplência é a Regressão Logística Múltipla. Segundo Rosa

(2000, p.18), trata-se de técnica usualmente empregada por profissionais da área financeira, pois apresenta um forte poder discriminante, não exige suposições fortes como na análise discriminante (variáveis independentes com distribuição normal multivariada), e seu algoritmo é facilmente encontrado em softwares estatísticos usuais, tais como o SPSS, usado para este trabalho. Em um modelo de Regressão Logística, a variável resposta ou variável dependente Y , assume valores dicotômicos:

$$Y = \begin{cases} 1, & \text{se contém o atributo;} \\ 0, & \text{em caso contrário.} \end{cases} \quad (1)$$

Nestas condições, a exemplo do modelo de regressão linear, a esperança condicional de Y dado um valor x , é expressa por $E(Y | x)$, (leia-se: esperança de Y dado o valor x), onde x é a variável independente. Em

(1), tem-se $E(Y=1 | x)$, se $Y=1$ ou, $E(Y=0 | x)$ se $Y=0$.

Para simplificar a notação, faça-se $E(Y | x) = \pi(x)$, o que representa a probabilidade de ocorrer o evento $\{y=1\}$ dado uma relação com a variável x .

O modelo a ser empregado, segundo Hosmer e Lemeshow (2000, p. 1-9), é apresentado na equação (2), a seguir demonstrada. Trata-se de uma Regressão Logística Múltipla, por ter uma quantidade k de

variáveis independentes. A variável dicotômica Y, conforme definido em (1), é função de um conjunto de variáveis independentes X_i , $i = 1, 2, \dots, k$.

$$\pi(X_i) = \frac{e^{\beta_0 + \sum_i \beta_i X_i}}{1 + e^{\beta_0 + \sum_i \beta_i X_i}} \quad (2)$$

Este modelo permite testar se o atributo dado à variável Y (classificação a priori), é função de k variáveis independentes x_k . Para tanto, é necessário estimar os coeficientes β_i , $i = 1, 2, \dots, k$, da equação (2), por meio da maximização da função de verossimilhança^[4] em relação aos k+1 elementos do vetor β . Para este trabalho, a estimativa dos coeficientes β_i e o cálculo do predito Y, foram realizadas com o uso do software estatístico *Statistical Package Social Science* (SPSS), versão 11.0.

2.1. A pesquisa realizada e os resultados obtidos

A busca pela resposta aos objetivos traçados torna este artigo um exercício empírico. Foi desenvolvido um modelo para mensurar o risco de inadimplência, com base em informações contábeis, usando os dados de clientes de uma instituição financeira nacional de grande porte. Foi selecionada uma amostra de 500 empresas, que apresentaram demonstrações contábeis e eram solicitantes e usuárias de crédito no período 2001 a 2002. A partir do estudo dos dados, construiu-se o modelo destinado a avaliar o risco de inadimplência.

O referencial teórico metodológico deste trabalho fundamenta-se nas classificações propostas por Abramo (1979, p. 34-44). É uma pesquisa aplicada, pois é empregada para que os resultados sejam usados na solução de problemas concretos, e estrutural, pois usa a análise sistemática da forma, do funcionamento, dos elementos e das suas inter-relações.

A coleta de dados exigiu a pesquisa em dois sistemas e respectivos arquivos de dados da instituição financeira:

- Sistema 1: dados cadastrais, constituído de balanços sintéticos de empresas solicitantes de crédito no referido banco;
- Sistema 2: dados de risco, com informações quanto ao tipo de empréstimo que o cliente solicitou, data de início e condições em que o crédito se encontra: em dia ou quando em atraso, por faixa de dias de atraso.

A estratégia de construção do plano de amostragem teve duas etapas:

- a) Coleta aleatória no arquivo do Sistema 1, de clientes que apresentassem o último balanço com data de dezembro de 2001, com um ou mais períodos de balanço anual, imediatamente anterior. O objetivo foi criar um arquivo contendo empresas com pelo menos dois balanços anuais,

para a construção dos índices contábeis. Estes índices seriam as variáveis independentes no modelo da Regressão Logística Múltipla. A retroatividade de amostra, de 1999 a 2001, ocorreu em função da necessidade de trabalhar com amostras mais atuais para a validação do modelo;

- b) Selecionar os clientes, e verificar quais se encontravam no arquivo do Sistema 2, durante 2002. A intenção nesta fase foi determinar a variável dependente Y, ao classificar *a priori* o cliente adimplente e o cliente inadimplente. Uma vez desenvolvido o modelo, é possível calcular a probabilidade de inadimplência para o período de um ano. Também nesta

etapa, foram selecionados os clientes tomadores de empréstimos de curto prazo, destinados às operações de curto prazo das empresas tomadoras do crédito: capital de giro, conta garantida e cheque especial empresarial.

Após a eliminação das duplicidades de clientes^[5], a amostra resultou em 500^[6] empresas de diversos ramos de atividades. Os clientes classificados na amostragem como adimplentes, foram aqueles que, durante 2002, apresentaram no máximo 60 dias de atraso. Clientes com atrasos superiores a 60 dias foram tratados como inadimplentes. A distribuição de frequência resultante é apresentada na tabela 1.

Tabela 1 – Classificação dos clientes pesquisados, segundo a situação de crédito:

Tipo de cliente	Frequência	%
Adimplentes	442	88,4
Inadimplentes	58	11,6
Total	500	100,0

Fonte: Elaborado pelos autores. São Paulo, 2015.

3. ÍNDICES FINANCEIROS USADOS NO MODELO

O estudo e a seleção dos índices financeiros baseou-se nas obras de Assaf

Neto (2002) e de Matarazzo (2003), Foram desenvolvidos e usados os índices a seguir apresentados no quadro 3:

Quadro 3 – Índices desenvolvidos e usados nas análises realizadas

i. ENDIVID: Participação de Capital de Terceiros:	i. PCCT: Composição do Endividamento
$\left(\frac{Pl - (Capital\ de\ Terceiros)}{Ativo\ Total} \right)$	$\frac{Passivo\ Circulante}{Capital\ de\ Terceiros} \times 100$
i. IMOBCP: Imobilização de Recursos não Correntes:	v. INDPFIN: Independência Financeira:
$\frac{Ativo\ Permanente}{PL + Exigível\ a\ Longo\ Prazo} \times 100$	$\frac{PL}{Ativo\ Total}$
v. LIQCORR: Liquidez corrente:	i. LIQSEC: Liquidez Seca:

$\frac{\text{Ativo Circulante}}{\text{Passivo Circulante}}$	$\frac{\text{Ativo Circulante} - \text{Estoque} - \text{Despesas Antecipadas}}{\text{Passivo Circulante}}$
i. LIQGER: Liquidez Geral:	i. VAT - Giro do ativo:
$\frac{\text{Ativo Circulante} + \text{Realizável a Longo Prazo}}{\text{Passivo Circulante} + \text{Exigível a Longo Prazo}}$	$\frac{\text{Vendas Líquidas}}{\text{Ativo}}$
k. LLV - Margem líquida:	k. LLAT - Rentabilidade do ativo:
$\frac{\text{Lucro Líquido}}{\text{Vendas Líquidas}} \times 100$	$\frac{\text{Lucro Líquido}}{\text{Ativo}} \times 100$
i. GAO - Alavancagem Operacional:	i. GAF - Alavancagem Financeira
$\frac{\text{Variação no Lucro Operacional}}{\text{Variação no Volume de Atividade}}$	$\frac{\text{Variação Percentual no Lucro Líquido}}{\text{Variação Percentual no Lucro Operacional}}$
i. GAT – Alavancagem Total:	
GAO x GAF	

Fonte: Elaborado pelos autores. São Paulo, 2015

O cálculo dos índices financeiros, segundo os critérios constantes no quadro 3, foi realizado com base nos balanços sintéticos usados pela instituição, relativo a 500

empresas componentes da amostra selecionada. Como primeiro resultado, a distribuição dos períodos pode ser verificada na tabela 2.

Tabela 2 - empresas segundo o ano de publicação do balanço

Anos de publicação	Empresas	%
1999 e 2000 e 2001	307	61,4
2000 e 2001	193	38,6
Total	500	100,0

Fonte: Elaborado pelos autores. São Paulo, 2015.

Em continuidade ao tratamento das variáveis independentes, os anos de publicação dos balanços, 1999, 2000 e 2001, foram categorizados, respectivamente como, períodos 1, 2 e 3. Desta forma, cada indicador

foi calculado para cada período. Calculou-se também, a variação percentual entre os índices, nos períodos 2 com 1, e 3 com 2. O Quadro 4 apresenta o resultado das relações:

Quadro 4 - períodos avaliados pelos índices definidos

Indicadores	Períodos	Indicadores	Períodos
1- ENDIVID- Participação de Capital de Terceiros	1, 2, 3, 2/1 e 3/2	8 - VAT - Giro do ativo	1, 2, 3, 2/1 e 3/2
2- PCCT – Composição do Endividamento	1, 2, 3, 2/1 e 3/2	9 - LLV - Margem líquida	1, 2, 3, 2/1 e 3/2
3 - IMOBCP - Imobilização dos Recursos não Correntes	1, 2, 3, 2/1 e 3/2	10 - LLAT - Rentabilidade do ativo	1, 2, 3, 2/1 e 3/2
4 - INDPFIN- Independência Financeira	1, 2, 3, 2/1 e 3/2	11 - GAO - Alavancagem Operacional	2/1 e 3/2
5 - LIQCORR - Liquidez corrente	1, 2, 3, 2/1 e 3/2	12 - GAF - Alavancagem Financeira	2/1 e 3/2
6 - LIQSEC - Liquidez Seca	1, 2, 3, 2/1 e 3/2	13 - GAT - Alavancagem Total	2/1 e 3/2
8- LIQGER - Liquidez Geral	1, 2, 3, 2/1 e 3/2		

Fonte: Elaborado pelos autores. São Paulo, 2015

É preciso observar que os valores dos índices são contínuos e, embora a Regressão Logística Múltipla permita o seu cálculo, a referida trabalha melhor com valores discretos, ou categorizados^[7]. Assim, como procedimento para categorização, calculou-se a estatística decil das respectivas distribuições. O valor do índice contábil referente ao i-ésimo decil, $i=1, 2, \dots, 9$; passou a ser o limite superior da classe da distribuição de frequência. No caso dos índices de ordem decrescente, após o cálculo do decil, reordenou-se a posição do valor associado ao decil de ordem i , $i=1, 2, \dots, 9$. Uma vez recodificadas, estas são as variáveis independentes a serem testadas no modelo de Regressão Logística.

3.1 RESULTADOS OBTIDOS

Para a escolha das variáveis que melhor explicassem a relação entre ser cliente adimplente e os índices financeiros abordados, aplicou-se o método *Stepwise*, teste *Forward*^[8], com probabilidade de entrada e saída iguais de 0,05 e 0,19 respectivamente. De acordo com Hosmer e Lemeshow (2000, p. 116), trata-se de um algoritmo estatístico de escolha de variáveis explicativas para o modelo, segundo uma regra fixa.

Sejam as medidas:

$$EF_1 = \frac{(Total\ de\ bons\ recomendados + total\ de\ maus\ não\ recomendados)}{Total\ Geral} \times 100$$

$$EM_1 = \frac{Total\ de\ maus\ não\ recomendados}{Total\ de\ maus} \times 100$$

$$EB_1 = \frac{Total\ de\ bons\ não\ recomendados}{Total\ de\ bons} \times 100$$

Fonte: Elaborado pelos autores. São Paulo, 2015

As probabilidades $\pi(x)$ calculadas no modelo expresso em (2), no ponto de corte

0,88^[9], conforme definido em EF_t , EB_t e EM_t estão dispostas na tabela 3:

Tabela 3 – percentagem de classificação correta

Passos do Stepwise	EF_t	EB_t	EM_t
1	70,4%	72,8 %	52,7%
2	68,2%	67,3%	74,5 %
3	70,6%	70,0 %	74,5 %

Fonte: Elaborado pelos autores. São Paulo, 2015

A eficiência total nas três simulações, no ponto de corte 0,88^[9] é próxima a 70 %; indicando uma estabilidade no modelo e um KS^[10] na ordem de 44,6 %. Neste estudo, optou-se pelos resultados do passo três. Porém, qualquer outra decisão poderia ser tomada.

Um teste importante fornecido pelo SPSS é a análise de resíduos: a diferença entre o valor observado e o valor estimado pelo modelo. A tabela 4 é o resultado da análise de resíduo no passo 3, denominado, Teste de Hosmer-Lemeshow.

Tabela 4 - dados para o teste de Hosmer-Lemeshow

Decil	Adimplentes		Inadimplentes		Total
	Observado	Estimado	Observado	Estimado	
1	25	27,49	22	19,51	47
2	36	34,49	10	11,51	46

3	40	38,13	6	7,87	46
4	39	39,63	6	5,38	45
5	42	42,09	4	3,91	46
6	44	43,34	2	2,66	46
7	44	44,11	2	1,89	46
8	45	44,68	1	1,32	46
9	44	45,28	2	0,72	46
10	45	44,76	0	0,24	45

Fonte: Elaborado pelos autores. São Paulo, 2015

O objetivo é testar as hipóteses:

H₀: Valor observado (Vobs) = Valor estimado (Vest)

H_a: Valor observado (Vobs) ≠ Valor estimado (Vest)

Para testar a hipótese nula, o artigo Logistic Regression (North Carolina State University, 2002, p. 8-9) fornece a seguinte regra de decisão:

A probabilidade *p-value* é obtida de uma distribuição de Qui-Quadrado com 8 graus de liberdade para teste da aderência do modelo. Se o Teste de Hosmer-Lemeshow é menor ou igual a 0,05, a hipótese nula H₀ é rejeitada; em caso contrário não é possível rejeitar a hipótese nula.

O SPSS forneceu a seguinte saída:

Quadro 5 – Teste de Hosmer-Lemeshow

Passos	Qui-Quadrado	Graus de liberdade	Significância
1	0,0000	8	1,000
2	6,926	8	0,545
3	4,243	8	0,835.

Fonte: Elaborado pelos autores. São Paulo, 2015

No passo três, não se rejeita H₀. Conclui-se, assim, que não é possível afirmar que os valores observados são significativamente diferentes dos valores estimados pelo modelo. Em consequência, as variáveis selecionadas no passo três são as

variáveis explicativas do modelo.

O modelo de Regressão Logística Múltipla, derivado da equação (2), está expresso na equação (3), a seguir apresentada:

$$\pi(X_i) = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3}} \quad (3)$$

Fonte: Elaborado pelos autores. São Paulo, 2015

Sendo: $\pi(x_i)$ é a probabilidade do i-ésimo cliente ser adimplente, $i=1, 2, \dots, 500$.

Com base no resultado do passo 3 do *Stepwise*, as variáveis selecionadas pelo modelo são:

- X₁: liquidez corrente no período 1 (anos 1999 e 2000);
- X₂: Variação no giro do ativo entre os períodos 2 e 1 (anos de 2001 e 2000);

- X₃: independência financeira no período 3 (anos 2000 e 2001).

O Quadro 6, por sua vez, apresenta as variáveis independentes escolhidas no passo 3, as respectivas estimativas dos parâmetros β (B), a Estatística de Wald, e Odds (exp(B)):

Quadro 6- variáveis selecionadas, correspondentes estimadores, Estatística de wald e ODDS

Variáveis	B	Wald	Exp(B)
Liquidez corrente (LIQCOR1)		19,68	
Até 0,5736	1,942	4,691	6,974
0,5736 -- 0,8369	1,302	2,631	3,676
0,8369 -- 1,0667	1,028	1,764	2,797
1,0667 -- 1,2648	1,753	3,983	5,771
1,2648 -- 1,5300	-0,420	0,365	0,657
1,5300 -- 1,8015	0,732	0,858	2,080
1,8015 -- 2,2136	0,170	0,056	1,186
2,2136 -- 3,3618	0,489	0,418	1,631
3,3618 -- 6,1514	2,282	3,981	9,793
Maior que 6,1514	0,000		1,000
Variação no giro de ativos V_AT21)		21,07	
Até -0,3641	-0,170	0,051	0,844
-0,3641 -- -0,2292	-0,876	1,477	0,417
-0,2292 -- -0,1355	1,101	1,334	3,007
-0,1355 -- -0,0793	0,232	0,084	1,261
-0,0793 -- -0,0257	-1,567	5,327	0,209
-0,0257 -- 0,0472	-0,271	0,115	0,763
0,0472 -- 0,1544	-1,145	2,563	0,318
0,1544 -- 0,3180	-0,166	0,047	0,847
0,3180 -- 0,6027	0,875	0,874	2,399
Maior que 0,6027	0,000		1,000
Independência financeira (NDEFIN3)		22,58	
Até 0,0525	-1,633	4,342	0,195
0,0525 -- 0,1616	0,134	0,027	1,143
0,1616 -- 0,2687	-0,264	0,113	0,768
0,2687 -- 0,3658	-0,489	0,183	1,399
0,3658 -- 0,4360	0,336	0,397	0,613
0,4360 -- 0,5248	-0,433	0,342	0,649
0,5248 -- 0,6143	0,815	0,853	2,259
0,6143 -- 0,7138	1,434	2,251	4,196
0,7138 -- 0,8175	2,013	3,024	7,482
Maior que 0,8175	0,000		1,000
Constante	1,716	4,826	5,564

Fonte: Elaborado pelos autores. São Paulo, 2015

A coluna B do Quadro 6, refere-se aos estimadores β do modelo. Em particular β_0 é a constante do modelo e no caso vale 1,716. Sua interpretação é similar aos modelos de regressão linear, e assim como ele, seu valor não tem significado lógico. A coluna Wald é a Estatística de Wald, uma medida para identificar qual, dentre as variáveis selecionadas, é a que mais participa na explicação da variável resposta. No modelo em questão, trata-se da variável Independência Financeira (INDEFIN3). Finalmente a coluna Exp (B), corresponde ao cálculo de e^{β} . Na literatura, esta expressão é conhecida como *Odds ratio*^[11].

3.2. TESTE DO MODELO PROPOSTO

Para validar o modelo (3), verificou-se o grau de acerto de sua previsão em recomendar ou não o crédito, com o real comportamento da empresa durante os meses de janeiro e fevereiro de 2003, quanto ao nível de atraso. O parâmetro de decisão do modelo foi o valor do ponto de corte 0,88, lembrando que esta é a proporção de clientes adimplentes na amostra de desenvolvimento.

A metodologia consiste em realizar a análise comparativa entre os resultados previstos com os reais obtidos. O valor 0,88 é comparado com a probabilidade $\pi(x)$ calculada em (3): se inferior a 0,88, tem o crédito recusado, em caso contrário, aprovado. Por outro lado, com referência aos meses de janeiro e fevereiro de 2003, verificou-se quais

das empresas já estudadas, na base de desenvolvimento, ainda estavam usando os créditos contratados em 2002. Nos casos em que ainda estivessem em vigor, após identificar o número de dias de atraso que a empresa apresentou, se superior a sessenta

dias, foi classificada como inadimplente, a exemplo da metodologia empregada na amostra de desenvolvimento.

A tabela 5 a seguir, apresenta os resultados do comportamento do modelo no período janeiro e fevereiro de 2003.

Tabela 5 - percentagem de classificação correta na amostra teste (jan/fev 2003)

A classificação correta na amostra teste (jun/rev 2006)			
Observado	Previsto		Kolmogorov – Smirnov (KS)
	Adimplentes	Inadimplentes	43,9%
Adimplentes	68,3%	24,4%	
Inadimplentes	31,7%	75,6%	
Eficiência do modelo	69,1%		

Fonte: Elaborado pelos autores. São Paulo, 2015

Da amostra inicial de 500 empresas analisadas na base de desenvolvimento, 375 (75%) apresentaram-se nos meses de janeiro e fevereiro de 2003. Dois motivos podem ser atribuídos à diferença: ou as empresas liquidaram a operação de crédito inicial e não a renovaram, ou liquidaram e contrataram um outro tipo de empréstimo que não fez parte deste estudo. Contudo, no conjunto, a estatística de Kolmogorov-Smirnov continuou boa, e o modelo manteve sua eficiência.

3.3 APLICAÇÃO DO MODELO

A equação (3) fornece a probabilidade de inadimplência para um ano, uma vez que o comportamento do cliente frente ao empréstimo foi observado durante o ano de 2002. Esta seção apresenta um método para indicar o valor de provisão de crédito, com base na classificação fornecida pela Resolução CMN nº 2.682/99. O método fundamenta-se na probabilidade fornecida pelo algoritmo apresentado na seção

anterior.

Estudos e recomendações do NBCA indicam que a análise do risco de crédito deve passar pela avaliação de três variáveis aleatórias: risco de inadimplência do cliente, risco de exposição do empréstimo (dado que o cliente se encontra inadimplente) e risco de não recuperar o crédito (dado que foi considerado perdido). Trata-se, assim, de avaliar respectivamente as probabilidades PD, EAD e LGD.

A metodologia apresentada a seguir supõe os casos extremos para os quais EAD e LGD são máximos^[12]. Dessa forma, o risco de crédito é o risco do cliente. Para classificar o risco do cliente em função da probabilidade $\pi(x_i)$ fornecida em (4), empregou-se $1-\pi(x_i)$: que é a probabilidade de inadimplência (*default*). A distribuição de perda revelou-se assintótica. O decil 9 revelou que 90% dos clientes apresentaram uma taxa de inadimplência inferior a 31%, conforme a tabela 6, a seguir apresentada.

Tabela 6 – Decis da distribuição de probabilidade de perda

Decil	Probabilidade de perda
1	0,0106
2	0,0224
3	0,0344
4	0,483
5	0,710
6	0,0958
7	0,1477
8	0,2078
9	0,3070
Total	481

Fonte: Elaborado pelos autores. São Paulo, 2015

Conhecendo-se a probabilidade de inadimplência, é possível estimar o número médio de empresas inadimplentes. Para tanto, foi necessário calcular a esperança da distribuição de inadimplência, expressa pela fórmula (5), a seguir apresentada.

$$E(\bar{\pi}) = E[1 - \pi(X)] = 58,38 \quad (5)$$

O valor 58,38 é a estimativa do número médio de clientes inadimplentes, estatisticamente igual ao número de clientes classificados como inadimplentes na amostra selecionada, conforme se verifica na tabela 1 (58 inadimplentes).

3.4 CLASSIFICAÇÃO DO RISCO DE INADIMPLÊNCIA E CONSEQUENTE PROVISÃO EM CRÉDITO

A metodologia para classificar o risco

não considerou a real distribuição de perda, pois como já visto, encontrava-se concentrada. Assim, optou-se por outro método:

- Classificar os clientes no nível de risco "AA", independentemente da probabilidade de inadimplência calculada pelo modelo. Definiu-se então que o número de empresas classificadas no risco "AA", limitar-se-ia ao valor do percentil 2 (P2), da distribuição de probabilidade de inadimplência. Sendo $P2 = 0,00289$, o que resultou em 10 empresas classificadas com risco "AA";

Os demais riscos, de um total de oito, foram obtidos segundo a função: $R_j = \frac{\sum_i^n (\bar{\pi}_i(X) f_i)}{\sum_i^n f_i} \leq k_j \%$,

sendo f_i a frequência de empresas cuja probabilidade de inadimplência é $\bar{\pi}_i(X)$; R_j = risco da classe j ; $j = 2, 3, 4, \dots, 9$; e K_j = é o

percentual máximo de provisão na classe de risco j , segundo o quadro a seguir:

Quadro 7 – Percentual de provisão máximo e risco correspondente

J	% de provisão (K)	Classe de risco	j	% de provisão (k)	Classe de risco
2	0,5	A	6	30,0	E
3	1,0	B	7	50,0	F
4	3,0	C	8	70,0	G
5	10,0	D	9	100,0	H

Fonte: Elaborado pelos autores. São Paulo, 2015

O resultado obtido com a aplicação do método encontra-se na

tabela 4, a seguir apresentada:

Tabela 7 - classe de risco para provisão em crédito

Classe de risco	% de provisão	Número de empresas	%	Provisionamento
AA	0,0	10	2,0	0,00
A	0,5	22	4,4	0,11
B	1,0	28	5,6	0,28
C	3,0	130	26,0	3,90
D	10,0	174	37,8	17,40
E	30,0	111	22,2	33,30
F	50,0	0	0,0	0,00
G	70,0	6	1,2	4,20
Total		481	100,0	59,19

Fonte: Elaborado pelos autores. São Paulo, 2015

Observando a tabela 7, verifica-se que não houve risco "F". Da classificação "E" passou-se imediatamente para o risco "G". O método de classificação apresenta uma concentração em níveis que exigem razoável provisão, mas na verdade, revelou-se muito adequado, uma vez que o número total de clientes a provisionar, 59, é bem próximo à proporção de clientes inadimplentes identificados na amostra de desenvolvimento (ver tabela 1), e na expectativa de inadimplência calculada em (5).

4 - CONSIDERAÇÕES FINAIS

Este trabalho apresenta um modelo estatístico para avaliar o risco de inadimplência. O modelo em questão permite aferir se o cliente é adimplente ou não em função de indicadores financeiros, obtidos nos demonstrativos das empresas selecionadas na amostragem. Os resultados obtidos permitem responder ao problema de pesquisa e corroborar a hipótese de trabalho, ambos apresentados no item 1.2 do texto.

O resultado do modelo é estatisticamente satisfatório. Os testes de validação conduzidos confirmaram a estabilidade e resultados próximos aos verificados na amostra de desenvolvimento. As informações obtidas nas demonstrações financeiras, utilizadas como insumo do modelo, propiciaram a condição de mensurar a probabilidade do cliente em cumprir ou não com o contrato de crédito.

Como consequência, e lembrando a abordagem sugerida pelo Novo Acordo da Basiléia no modelo IRB avançado, o qual define que o risco de crédito é função do risco de inadimplência do cliente (PD), da exposição do empréstimo tomado (EAD) e da taxa de perda em créditos inadimplentes (LGD), o modelo expresso em (3) atende ao cálculo de Probabilidade de Perda, como também às disposições regulatórias da Resolução CMN nº 2.682/99.

NOTAS DE FIM DE TEXTO

[1] Segundo o BANK FOR INTERNATIONAL SETTLEMENTS, *The New Basel Capital Accord Bank* (2003, p.6) a implementação do modelo (IRB) já deverá iniciar-se em 2007.

[2] *Rating*, classificação fornecida por empresas especializadas na prestação de informações. Avaliam preliminarmente a situação econômico-financeira do tomador, e por meio de classificação, revelam a capacidade da empresa em tomar crédito. Modelos de *rating* fornecem parâmetros de referência, com os quais os agentes do mercado se baseiam durante suas atividades.

[3] Ver, por exemplo, Saunders (2000), Securato (2002).

[4] Sobre Função de Verossimilhança ver, por exemplo, De Groot (1975).

[5] Na ocorrência de um cliente com mais de um empréstimo válido para o estudo, selecionou-se aquele que apresentasse empréstimo em pior situação de atraso.

[6] Supondo uma proporção máxima de clientes adimplentes e inadimplentes em 50%, e um nível de significância de 5%, a margem de erro para uma amostra de tamanho $n = 500$ é de $\pm 4,38\%$.

[7] Segundo Rosa (2000, p.14), as vantagens para categorizar as variáveis são: padronizar resultados, estabilizar o modelo e transformar as variáveis.

[8] Ver por exemplo, MENDENHALL; SINCICH (1996) ou consultar SPSS (1997).

[9] O ponto de corte separa os clientes recomendados pelo modelo, daqueles não recomendados. No caso aplicou-se a mesma proporção verificada na amostra selecionada.

[10] A manutenção deste valor tem como objetivo verificar se as condições da amostra de desenvolvimento se mantêm. O mercado financeiro ajusta o ponto de corte às necessidades da política de crédito, ou então pode empregar a estatística de Kolmogorov-Smirnov (KS), para escolher o valor que mais discrimina clientes adimplentes dos inadimplentes, entre outros métodos. Sobre a estatística de Kolmogorov-Smirnov consultar, por exemplo, Oliveira; Andrade (2002).

[11] "Odds of an event occurring are defined as the ratio of the probability that it will occur to the probability that it will not". (SPSS, 1997, p.42). Tradução livre: Odds de uma ocorrência de um evento são definidos como a razão da probabilidade em que ele ocorrerá, pela probabilidade de que ele não ocorrerá.

[12] Seria a situação em que o EAD e LGD tenham taxa igual a 100%. Ou seja, a exposição do empréstimo (dado que o cliente está inadimplente e as taxas de recuperação são respectivamente 100% e zero%). Nesse caso, $P(EAD)=P(LGD)=1$ e o risco de crédito é o risco do cliente, isto é PD.

REFERÊNCIAS BIBLIOGRÁFICAS E ELETRÔNICAS

ABRAMO, Perseu. Pesquisa em ciências sociais. In Sedi Hirano (coord.). Pesquisa social: projeto e planejamento. São Paulo: T. A. Queiroz, Editor, 1979. p.21-88.

ASSAF NETO, Alexandre. Estrutura e análise de balanços: um enfoque econômico financeiro. 7ªed. São Paulo: Atlas, 2002. 320 p.

BANK OF INTERNATIONAL SETTLEMENTS. Consultative document overview of the new Basel Capital Accord: Basel committee on banking supervision. Basel: april 2003. Versão disponível em: (<http://www.bis.org>).

COMITÊ DE SUPERVISÃO BANCÁRIA DA BASILÉIA. O método baseado em classificações internas. Tradução: Ernest & Young, 2001. versão 1.0. Texto disponível em: <<http://www.febraban.com.br>> Acesso em: 25 jul. 2001. Versão original em inglês disponível em: (<http://www.bis.org>).

_____. Visão geral do novo Acordo de capital da Basileia. Tradução: Ernest & Young, 2001. versão 1.0. Texto disponível em: <<http://www.febraban.com.br>> Acesso em: 25 jul. 2001. Versão original em inglês disponível em: (<http://www.bis.org>).

CONSELHO MONETÁRIO NACIONAL. Resolução nº. 2.682, de 21 de dezembro de 1999. Texto disponível em: <<http://www.bcb.gov.br>> Acesso em: 25 jun. 2001.

DE GROOT, Morris H. Probability and statistics. 2ªed. United States of America: Addison -Wesley Publishing Company, 1975. 723p.

HOSMER, David W.; LEMESHOW, Stanley. Applied logistic regression. 2ªed. New York: A Wiley – Interscience Publication, 2000. 373p.

MATARAZZO, Dante.C. Análise financeira de balanços. 6ªed. São Paulo: Atlas, 2003. 459 p.

MENDENHALL, William; SINCICH, Terry. A second course in statistics: regression analysis. 5. ed. New Jersey: Prentice-Hall, Inc. 1996. 899p.

NORTH CAROLINA STATE UNIVERSITY. Logistic regression. Texto disponível em: <<http://www.2.chass.ncsu.edu/garson/pa765/index.html>> Acesso em 15 jun. 2003.

OLIVEIRA, José G. C.; ANDRADE, Fábio W. M. Comparação entre medidas de performance de modelos de credit scoring. Revista de Tecnologia de Crédito. São Paulo: Serasa, VI, no. 33, set. 2002. 75p.

ROSA, Paulo T. M. Modelos de "credit score": regressão logística, Chaid e Real. 2000. 69f. Dissertação (Mestrado em Estatística). Instituto de Matemática e Estatística da Universidade de São Paulo. São Paulo, 2000.

SAUNDERS, Anthony. Medindo o risco de crédito. Tradução: Bazán Tecnologia e Linguística Ltda. Rio de Janeiro: Qualitymark, 2000a. 181p.

SECURATO, José R. (coord.). Crédito: análise e avaliação do risco. São Paulo: Saint Paul Institute of Finance, 2002. 355p.

SPSS: Professional Statistics™ 7.5. Chicago: SPSS Inc., 1997.

