

ÁREA TEMÁTICA: FINANÇAS E ECONOMIA

# Gestão de Universidade Privada: risco de inadimplência dos alunos

*Gestión de la Universidad Privada: el riesgo del impago de los alumnos*

*Private University Management: default risk by students*

Cristiane Freitas Ribeiro<sup>1</sup>

João Zani<sup>2</sup>

Clea Beatriz Macagnan<sup>3</sup>

Francisco Antonio Mesquita Zanini<sup>4</sup>

Recebido em 17 de julho de 2011 / Aprovado em 28 de março de 2012

Editor Responsável: João Maurício Gama Boaventura, Dr.

Processo de Avaliação: *Double Blind Review*

## RESUMO

Este artigo analisa a probabilidade do risco de inadimplência em uma universidade privada (UP). Nele, utilizou-se a técnica estatística de regressão logística. O modelo de risco de crédito foi construído com base em uma amostra de alunos matriculados em uma UP, situada no Rio Grande do Sul. As variáveis explicativas do modelo foram obtidas a partir da aplicação de um questionário socioeconômico, que gerou um total de 59 variáveis, das quais apenas três foram representativas: existência de débitos já negociados, posse de cartão de crédito e reprovação. Os resultados obtidos mostram que o modelo estudado obtém resultados satisfatórios, quando aplicado na medição da probabilidade de risco de crédito.

**Palavras-chave:** Universidade privada. Risco. Inadimplência.

## ABSTRACT

This paper analyzes the probability of default risk in a private university (UP) using the statistical technique of logistic regression. The credit risk model used was based on a sample of registered UP students in Rio Grande do Sul. The explanatory variables of the model were obtained by applying a socioeconomic questionnaire that generated a total of 59 variables, of which only three were representative: the existence of debts that had already been negotiated, credit card ownership and rejection. The obtained results show that the

1. Mestre em Ciências Contábeis pela Universidade do Vale do Rio dos Sinos – UNISINOS. Professora na Universidade Católica de Pelotas – UCPEL. [tianefer@gmail.com]

2. Doutor em Administração pela Universidade Federal do Rio Grande do Sul – UFRGS. Professor na Universidade do Vale do Rio dos Sinos – UNISINOS. [jzani@unisinis.br]

3. Doutora em Creación, Estrategia y Gestión de Empresas pela Universidad Autonoma de Barcelona, Espanha. Professora na Universidade do Vale do Rio dos Sinos – UNISINOS. [cleabeatriz@hotmail.com]

4. Doutor em Contabilidade e Organização de Empresas pela Universidad Autonoma de Madrid, Espanha. Professor na Universidade do Vale do Rio dos Sinos – UNISINOS. [fzanini@unisinis.br]

Endereço dos autores: Av. Unisinis, 950 – Cristo Rei, São Leopoldo – RS Cep. 93022-000 Brasil

studied model yields satisfactory results when used to measure credit risk probability.

**Key words:** Private university. Risc. Default.

## RESUMEN

El artículo analiza la probabilidad del riesgo de los estudiantes no realizaren el pago de la deuda en una universidad privada (UP). Se ha utilizado la técnica estadística de regresión logística. El modelo de riesgo de crédito ha sido construido con base en una muestra de estudiantes matriculados en la respectiva UP, ubicada en el Estado del Rio Grande do Sul. Las variables explicativas del modelo han sido obtenidas a partir de la aplicación de un cuestionario socioeconómico, que ha generado el total de 59 variables de las cuales solamente tres han sido representativas: existencia de débitos ya negociados, propiedad de tarjetas de crédito y reprobación. Los resultados obtenidos muestran que el modelo estudiado tiene resultados satisfactorios, cuando aplicado en la medición de probabilidades de riesgo de crédito.

**Palabras clave:** Universidad privat. Riesgo. Incumplimiento contractual.

## 1 INTRODUÇÃO

A consolidação do ambiente macroeconômico mais favorável no Brasil, oportunizado também pelo crescimento da economia mundial, possibilitou uma diminuição importante nas taxas de juros no país e consolidou fatores condicionantes para a sustentabilidade do nosso crescimento econômico nos últimos cinco anos. Esses fatores soaram como um estímulo à flexibilização das condições de prazo e taxas das operações de crédito, tanto no âmbito das pessoas jurídicas como no das pessoas físicas. As operações de crédito do Sistema Financeiro Nacional cresceram 20,7% no ano de 2006.

Analizando os dados a partir da distribuição do crédito por segmento de atividade econômica, destacam-se os setores relacionados a empréstimos às pessoas físicas, à indústria e ao comércio como os

de maior expansão. Somente no segmento de linhas de crédito, destinadas às pessoas físicas, o acréscimo correspondeu a 24,9%, entre os anos de 2005 e 2006, destinados, especificamente, a contratos vinculados à aquisição de veículos e financiamentos para crédito pessoal. Da mesma forma, assim como ocorreu o crescimento da política de concessão de crédito, ocorreu o crescimento dos níveis de inadimplência relacionados a essas operações. Em 2006, o índice de inadimplência de operações com pessoas físicas elevou-se em 0,9 p.p., traduzindo um aumento do saldo dos créditos inadimplentes de 20,9%, quando comparado ao ano de 2005 (BANCO CENTRAL DO BRASIL, 2006).

Esta combinação gerou: aumento das operações de crédito *versus* aumento dos níveis de inadimplência, impulsionando um movimento das instituições, no sentido de programar melhores sistemas de gerenciamento de crédito. Na prática, busca-se uma ferramenta mais robusta, segura e confiável que traduza as informações sobre os clientes tomadores de crédito. Esta realidade configura-se no segmento de atividade econômica das universidades privadas (UP). As razões para isso ocorrer estão relacionadas a diversos elementos; entre eles destaca-se a facilidade existente na criação dessas UPs, que se traduz em um crescimento exacerbado, frágil planejamento do setor. Ainda, pelo aumento de alunos concluintes do ensino médio, que resulta em maior demanda de vagas nas universidades, incluindo a população com menor renda. O que gera aumento na busca pela concessão de crédito proveniente de programas de Crédito Educativo, patrocinados pelo governo federal (Fies e ProUni) e de bolsas de estudos ou outras modalidades de financiamento (LOCH; REIS, 2004, p. 1).

A correlação estabelecida aqui é que a majoração do número de UPs e, conseqüentemente, a elevação do número de alunos de baixo poder aquisitivo com acesso a essas instituições denota a ideia de crescimento das operações de crédito destinadas a esse segmento. Similarmente às instituições financeiras de crédito, as instituições de ensino superior buscam outras formas de gerenciar o risco de inadimplência, além dos tradicionais qualitativos, como servir-se de dados do

cadastro no Serviço de Proteção ao Crédito (SPC) e a cobrança administrativa e judicial dos inadimplentes. A procura passa a ser por mecanismos mais inteligentes, que concedam subsídios para analisar o processo, desde a concessão do crédito até o momento de sua liquidação.

Neste contexto, o objetivo deste estudo é analisar a probabilidade do risco de inadimplência em uma UP. A partir da construção do modelo, desenvolve-se análise do perfil dos usuários, no caso alunos da universidade que, então, assumem o papel de tomadores de crédito, a partir de variáveis como renda, estado civil, número de filhos, entre outras. Igualmente, é objetivo identificar as variáveis capazes de explicar o modelo e validá-lo pela classificação dos alunos dentro da probabilidade de tornarem-se adimplentes ou inadimplentes.

## 2 REVISÃO DA LITERATURA

Crédito significa crença, confiança ou fé. Segundo Securato (2002, p. 17), “o termo crédito estabelece uma relação de confiança entre duas (ou mais) partes em uma determinada operação”. Neste sentido, o crédito produz dois efeitos, o de acreditar na promessa de pagamento e no intervalo estabelecido entre a aquisição e a liquidação da dívida. Muito embora a confiança seja um elemento importante nas relações contratuais, não seria suficiente para uma concessão de crédito. Estabelecer um contrato de crédito com base apenas na confiança, sem estabelecer as garantias necessárias para diminuir o risco de inadimplência do tomador, poderia significar um comportamento bastante ingênuo.

Parte-se do pressuposto de que o contrato é a representação de um acordo, em que as partes envolvidas comprometem-se mutuamente, no futuro determinado, a agir conforme comprometido. Do comportamento futuro, são conhecidas apenas as probabilidades e não a certeza. Assim, haveria sempre um determinado grau de risco nas decisões de concessão de crédito, o qual se caracteriza pelo risco de uma contraparte, em um acordo de concessão de crédito, não honrar seu compromisso (GONÇALVES, 2005, p. 22). Para miti-

gação do risco, são utilizadas basicamente duas formas ou técnicas de análise:

- qualitativa: quando a concessão do crédito é estabelecida a partir de uma avaliação do analista de crédito, marcada pela subjetividade.
- quantitativa: quando há o uso de métodos estatísticos.

A técnica qualitativa é a mais tradicional, baseia-se em informações cadastrais, documentos comprobatórios (comprovantes de renda e demonstrações financeiras), históricos de operações de crédito passadas e perspectivas futuras de desempenho. Sua característica principal é a subjetividade, em função de que a análise do risco de crédito que, embora seja efetuada por profissionais treinados, depende da percepção individual do analista. A técnica quantitativa é complementar às evidências obtidas a partir dos métodos tradicionais, visa possibilitar mais segurança, estruturação e gerenciamento do processo de tomada de decisão da concessão do crédito.

### 2.1 Técnicas de avaliação do risco de crédito

A avaliação do risco de crédito possui duas vertentes: a avaliação de risco para pessoas jurídicas e a avaliação de risco para pessoas físicas. Em ambos os casos, são aplicadas técnicas estatísticas com o intuito de melhor garantir a securitização do processo. Com relação às técnicas estatísticas de avaliação de risco de crédito à pessoa jurídica, os estudos relacionam-se à previsão da probabilidade de solvência ou falência das empresas. Entre as técnicas mais utilizadas, destacam-se a análise discriminante linear e a regressão logística. Para concessão de crédito para pessoa física, temos dois modelos conceituais: o *credit scoring* e o *behavioural scoring* (THOMAS, 2000, p. 149-172). Entre as técnicas estatísticas utilizadas nos modelos de *credit scoring*, pode-se citar: a análise discriminante e a regressão logística.

A análise discriminante “é uma técnica estatística multivariada usada na resolução de pro-

blemas que envolvem a separação de conjuntos distintos de observações e a alocação de novas observações em um conjunto específico”. Integra o conjunto de técnicas usadas no reconhecimento de padrões (GUIMARÃES; CHAVES NETO, 2002, p. 4). Ela é uma das técnicas estatísticas mais empregadas para determinar o peso dos índices em modelos de *credit scoring*. 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” (BRITO; ASSAF NETO, 2008, p. 23). Nesses casos, quando a variável dependente é definida pela escolha entre duas opções, diz-se que ela deixa de ser quantitativa e passa a se tornar qualitativa. Para estes modelos de regressão, o objetivo da análise é encontrar a probabilidade de que algo aconteça (GUJARATI, 2006, p. 470) e, por conseguinte, são denominados de modelos de probabilidade. Na sequência apresenta-se o quadro 1, com características de cada técnica.

O *behavioural scoring* é “um sistema de pontuação com base em análise comportamental das informações que a empresa já possui sobre o cliente na renovação ou concessão de uma nova linha de crédito” (QUEIROZ, 2006, p. 5). Diferentemente do *credit scoring*, o *behavioural scoring* analisa o comportamento passado do tomador do recurso. Verifica como o indivíduo usou o crédito, quanto usou, o que usou, se mostra qualquer indicativo de atrasos e outras informações dispo-

níveis (MINUSSI, 2001). As informações adicionais extraídas desse sistema são o reembolso e a história de requerimentos de crédito deste consumidor (THOMAS, 2000).

### 3 METODOLOGIA

#### 3.1 Definição da amostra

Para a construção do modelo, é necessário obter uma amostra constituída de indivíduos classificados como adimplentes e inadimplentes, o qual define que a variável dependente deve ser categórica e deve assumir um entre dois resultados possíveis (binária), tais como: “adimplentes ou inadimplentes” e “bom pagador ou mau pagador” (BRITO; ASSAF NETO, 2008, p. 23).

Foram considerados inadimplentes, os alunos com “falta de cumprimento das cláusulas contratuais em determinado prazo” (SANDRONI, 2006, p. 416). A política de crédito adotada pela IPES (Instituição Privada de Ensino Superior), a qual os alunos entrevistados se encontram matriculados, permite que cada um efetue sua matrícula a cada início de semestre e divida o pagamento em seis parcelas subsequentes.

Em outras palavras, um aluno cuja matrícula foi efetivada em dezembro de 2007 para cursar disciplinas no 1º semestre de 2008 pode parcelar o valor correspondente em seis parcelas vencíveis entre janeiro e junho de 2008. Para a IPES, inadimplente é o aluno que tem parcela em atraso de prazo igual ou superior a 45 dias. A

Análise Discriminante	Regressão Logística
Normalidade das variáveis independentes	Distribuições entre as variáveis não normais
Interpretação por meio de escores	Interpretação por meio de probabilidade
Variável dependente não métrica	Variável dependente categórica
Matrizes de variância e covariância iguais	Relação não linear entre a variável dependente e as variáveis independentes

**Quadro 1** – Comparação entre os métodos de *credit scoring* utilizados em análise de risco de crédito pessoa física.

Fonte: dos autores.

partir desta definição, ao final do mês de junho de 2008, efetuou-se o levantamento da relação de todos os alunos efetivamente matriculados no semestre que possuíam débitos atrasados em prazos iguais ou superiores a 45 dias, obtendo-se uma amostra inicial de 793 alunos, de um total de matriculados na universidade, no mesmo período, igual a 5.386.

Ao comparar o total de alunos matriculados na IPES no 1º semestre de 2008 (5.386 alunos) com o total de alunos inadimplentes (793 alunos), verificou-se que o montante de alunos inadimplentes corresponde a 14,72% da população total.

Ao final do mês de junho de 2008, a pesquisa atingiu um montante de 392 alunos. Aplicando-se os critérios pré-estabelecidos pela própria IPES, que definia o aluno adimplente e o aluno inadimplente e, conjuntamente, a avaliação de que todas as observações estavam em perfeitas condições de análise, a amostra final ficou dimensionada em 206 alunos. Desses, 32 alunos apresentavam a condição de falta de pagamento em alguma de suas parcelas, com prazo igual ou superior a 45 dias. Nessa condição, o vínculo estabelecido entre o montante de alunos inadimplentes (32 alunos) e o tamanho da amostra dimensionada (206 alunos) equivale a 15,53%.

Comparando-se a população-alvo do estudo com a amostra de trabalho obtida, verifica-se que os percentuais de alunos inadimplentes correspondem a aproximadamente 15%, o que valida a utilização da amostra, visto estar de acordo com a realidade apresentada pela instituição. A composição da amostra aleatória final foi de 206 alunos, qualificada por dois grupos de estudos específicos: adimplentes e inadimplentes, compostos, respectivamente, por 85% e 15% dos indivíduos, percentuais correspondentes à verdadeira situação encontrada na instituição.

### 3.2 Instrumento de obtenção dos dados

O instrumento de coleta utilizado para a obtenção das evidências que compõe as variáveis capazes de explicar a situação de adimplência e inadimplência dos alunos foi um questionário

socioeconômico (Apêndice A). O questionário foi testado a partir de um pré-teste, conforme recomendado por Hair Junior *et al.* (2005, p. 30), em que se avaliou a exatidão e coerência das respostas, uma vez que foi aplicado a uma pequena amostra de respondentes com características semelhantes à população-alvo.

O pré-teste foi feito nas salas de aulas, com uma amostra de 58 alunos. Após o processo de validação do questionário, pôde-se estender a sua aplicação à população-alvo do estudo: 5.386 alunos.

O instrumento foi disponibilizado aos alunos por meio do ambiente virtual corporativo da IPES e enviada uma mensagem ao e-mail particular de cada aluno. Muito embora a utilização da internet possibilite ao pesquisador agilidade no processo, baixo custo, proporcionando uma alimentação on-line e automatizada do banco de dados, e uma redução do viés existente entre o entrevistador e o respondente (FREITAS *et al.*, 2004, p. 3), não se pode deixar de destacar seus pontos negativos, que devem ser observados quando da aplicabilidade deste recurso.

Caberia destacar: falta de segurança da internet como meio de coleta, já que este instrumento ainda gera desconfiança, resistência e ansiedade em seus respondentes, problemas de interpretação em razão de não haver interação direta com o aplicador, impossibilidade dos respondentes acessarem a internet ou o computador, a falta de familiaridade com esta ferramenta e a qualidade técnica insuficiente, mesmo quando o instrumento é validado e padronizado (JOLY; SILVEIRA, 2003, p. 3).

Por meio da aplicação do questionário, foi obtido um total de 392 respondentes, entre os quais foram selecionados 206 alunos, amostra do estudo, sendo 174 alunos adimplentes e 32 alunos inadimplentes.

### 3.3 Variáveis explicativas

Após a coleta e tabulação das evidências obtidas por meio do instrumento selecionado, neste caso um questionário socioeconômico, adotaram-se, inicialmente, as seguintes variáveis expli-

cativas (Quadro 2) para constituir o modelo econométrico para a estimação da função logística de probabilidade do risco de crédito:

As variáveis X1, X2, X3, X4, X5, X6, X7, X8, X9, X10, X11, X12, X13, X14, X15 e X16 referem-se às informações pessoais do aluno. As variáveis X17, X18, X19, X20 e X21, X22, X23, X24, X25, X26, X27, X28, X29, X30, X31 e X32 referem-se às informações acadêmicas e financeiras do aluno para com a uni-

versidade. As variáveis X33, X34, X35 e X36 referem-se a informações profissionais do aluno. As variáveis X37, X38, X39, X40, X41, X42, X43, X44, X45, X46, X47, X48, X49, X50 e X51 referem-se às informações do grupo familiar do aluno. As variáveis X52, X53, X54, X55, X56, X57 e X58 referem-se às responsabilidades financeiras do aluno e a variável X59 refere-se à informação sobre a instituição privada de ensino superior.

Variáveis explicativas	
X1: sexo	X31: já negociou débitos até duas vezes
X2: idade abaixo de 20 anos	X32: já negociou débitos acima de duas vezes
X3: idade de 20 a 25 anos	X33: possui atividade remunerada
X4: idade mais de 25 anos	X34: atividade remunerada de até R\$ 415
X5: naturalidade	X35: atividade remunerada acima de R\$ 415 a R\$ 4.150
X6: casa própria	X36: atividade remunerada mais de R\$ 4.150
X7: reside com família	X37: pai sem escolaridade
X8: reside com parentes	X38: pai com ensino fundamental/médio
X9: reside com outros	X39: pai com ensino superior/pós-graduação
X10: solteiro	X40: mãe sem escolaridade
X11: casado	X41: mãe com ensino fundamental/médio
X12: outro estado civil	X42: mãe com ensino superior/pós-graduação
X13: filhos ou dependentes	X43: existem dependentes da renda
X14: de 0 a 1 dependente	X44: até dois dependentes
X15: mais de 1 a 3 dependentes	X45: mais de dois dependentes
X16: mais de 3 dependentes	X46: renda familiar até R\$ 415
X17: de 3 a 10 disciplinas cursadas	X47: renda familiar mais de R\$ 415 a R\$ 4.150
X18: de 10 a 20 disciplinas cursadas	X48: renda familiar mais de R\$ 4.150
X19: mais de 20 disciplinas cursadas	X49: possui familiar no ensino superior
X20: de 3 a 10 disciplinas a cursar	X50: até dois familiares no ensino superior privado
X21: de 10 a 20 disciplinas a cursar	X51: acima de dois familiares no ensino superior privado
X22: mais de 20 disciplinas a cursar	X52: possui financiamento
X23: reprovação	X53: financiamento de até R\$ 415
X24: reprovação até 2 vezes	X54: financiamento acima de R\$ 415 até R\$ 4.150
X25: reprovação mais de 2 vezes	X55: financiamento acima de R\$ 4.150
X26: mensalidade em atraso	X56: possui conta bancária
X27: débito até mil reais	X57: possui cartão de crédito
X28: débito de mil a cinco mil reais	X58: possui seguro de vida
X29: débito acima de 5 mil reais	X59: qualidade do ensino
X30: já negociou débitos	

**Quadro 2** – Variáveis explicativas iniciais.

Fonte: dos autores.

### 3.4 Desenvolvimento do modelo

Para o desenvolvimento deste estudo, optou-se pelo uso do método *credit scoring*. Entre as técnicas estatísticas utilizadas em sistemas de *credit scoring*, optou-se pela técnica de regressão logística. Isso porque estudos como de Minussi (2001), Brito e Neto (2008) obtiveram bons resultados com a utilização do modelo *logit*.

Para o desenvolvimento do modelo econômico, proposto por este estudo, era necessário estabelecer algumas relações entre a variável dependente e as variáveis explicativas, objetivando identificar quais variáveis são capazes de estimar a probabilidade do risco de inadimplência. Para uma verificação estatística de dependência entre a variável a ser explicada ( $Y$ ), probabilidade do risco de crédito e cada variável explicativa ( $X_i$ ), foi aplicado o teste não paramétrico do qui-quadrado.

Esse teste foi adotado em razão de as variáveis serem dados categorizados, possibilitando ser aplicado independentemente da forma da distribuição da população da qual se extraiu a amostra aleatória (SIEGEL, 1975, p. 35). A Tabela 1 apresenta os resultados do teste de independência estatística para as variáveis regressoras/explicativas:

Conforme é possível observar na Tabela 1, fica confirmada a significância das variáveis X23, X26, X30, X31, X32, X49, X52 e X57 com a variável  $Y$ , informando que existe relação de dependência com a inadimplência. Entre as várias suposições estabelecidas para a técnica estatística de regressão logística, uma das mais importantes e que pode causar preocupação e distorção na análise da significância dos coeficientes obtidos é a figura da colinearidade ou multicolinearidade. A análise da multicolinearidade verifica se há correlação entre as variáveis regressoras  $X_i$ , dificultando a separação dos efeitos de cada uma delas sobre a variável dependente  $Y$  (CORRAR; PAULO; DIAS FILHO, 2007, p. 156). O teste aplicado para a análise de multicolinearidade entre os  $X_i$  significativos foi o de Correlação de Spearman e seus resultados estão apresentados na Tabela 2.

Embora os resultados do teste de correlação tenham selecionado quatro variáveis regressoras possíveis de serem aplicadas, os testes de avaliação do ajuste do modelo não confirmam a mesma hipótese. Neles, a variável qualidade de ensino é considerada não significativa aos níveis de 1% e 5%.

Sendo assim, o modelo final de regressão logística foi composto pelo conjunto de três das 59 variáveis explicativas que faziam parte da análise.

**Tabela 1** – Resultado do Teste Qui-Quadrado para as variáveis regressoras significativas.

Descrição	Inadimplência		
	Sign.	X <sup>2</sup>	P
X23: reprovação	***	7,6	0,050
X26: mensalidade em atraso	**	329,2	0,000
X30: já negociou débitos	**	329,2	0,000
X31 e X32: quantas vezes negociou débitos	**	9,57	0,001
X49: possui familiar no ensino superior	**	7,34	0,006
X52: possui financiamento	***	4,41	0,030
X57: possui cartão de crédito	**	7,58	0,000

Fonte: dos autores

(\*\*) Estatisticamente significativo a um nível de 1%

(\*\*\*) Estatisticamente significativo a um nível de 5%

Tabela 2 – Resultados do teste de correlação de Spearman para a colinearidade.

Descrição	Reprovação		Já negociou débitos		Possui cartão crédito		Qualidade do ensino	
	S	Resultado	S	Resultado	S	Resultado	S	Resultado
Qualidade de ensino	0,042	N/Signif.	-0,074	N/Signif.	-0,015	N/Signif.	-	-
Reprovação	-	-	0,165	N/Signif.	-0,040	N/Signif.	0,042	N/Signif.
Já negociou débitos	0,165	N/Signif.	-	-	-0,005	N/Signif.	-0,074	N/Signif.
Possui cartão crédito	-0,040	N/Signif.	-0,005	N/Signif.	-	-	-0,015	N/Signif.

Fonte: dos autores

(\*) Estatisticamente não significativo

(\*\*) Estatisticamente significativo com 5%

(\*\*\*) Estatisticamente significativo com 1%

As variáveis incluídas no modelo foram X<sub>23</sub> (reprovação), X<sub>30</sub> (já negociou débitos) e X<sub>57</sub> (possui cartão de crédito). Sendo assim, a função matemática correspondente ao modelo desenvolvido é a Equação (1), a seguir:

$$\ln\left(\frac{p}{1-p}\right) = -2,457 + 0,616 X_{23} + 2,196 X_{30} - 1,263 X_{57} \quad (1)$$

Onde,

X<sub>23</sub> = Reprovação

X<sub>30</sub> = Já negociou débitos

X<sub>57</sub> = Possui cartão de crédito

Da mesma forma, o modelo de estimação da função risco de crédito obtido e de melhor ajuste aos dados empíricos é dado pela Equação (2):

$$P = \frac{1}{1 + e^{-[-2,457 + 0,616 X_{23} + 2,196 X_{30} - 1,263 X_{57}]}} \quad (2)$$

Onde,

P = Representa a probabilidade do risco de inadimplência do aluno

X<sub>23</sub> = Reprovação

X<sub>30</sub> = Já negociou débitos

X<sub>57</sub> = Possui cartão de crédito

Analisando o modelo, verifica-se que os coeficientes das variáveis apresentaram o sinal esperado na função logística. As variáveis com coeficientes positivos indicam que, quanto maiores forem seus valores, maior será a probabilidade do risco de inadimplência do aluno. A variável

com coeficiente negativo indica que, quanto maior for o seu valor, menor será a probabilidade do risco de inadimplência do aluno.

A explicação para isto é a probabilidade de ocorrência do risco de inadimplência estar relacionada ao aluno que já sofreu alguma reprovação (coeficiente com sinal positivo), já negociou algum débito com a instituição (coeficiente com sinal positivo) e não possua cartão de crédito (coeficiente com sinal negativo).

### 3.5 Avaliação do ajuste do modelo

Para o modelo logístico estimado foram realizados vários testes estatísticos que validam a sua aplicabilidade. A intenção aqui é informar os testes realizados na avaliação da qualidade do modelo de regressão logística aplicado neste estudo. O programa computacional utilizado para a geração dos dados estatísticos foi o SPSS 13.0. Em regressão logística, o método utilizado para estimar os coeficientes é o de máxima verossimilhança; ou seja, “ao invés de minimizar os desvios quadráticos (mínimos quadrados), a regressão logística maximiza a “verossimilhança” de que um evento ocorra” (HAIR JUNIOR *et al.*, 2005, p. 234).

Para avaliar a significância dos coeficientes da regressão logística foi aplicada a estatística Wald. O Teste Wald avalia a hipótese nula de o parâmetro estimado ser igual a zero (BRITO; ASSAF NETO, 2008, p. 26), o que significa que a probabilidade

**Tabela 3** – Modelos estimados pelo teste estatístico de regressão logística.

Modelos/Variáveis	$\beta_1$	Wald	$\beta_2$	Wald	$\beta_3$	Wald	$\beta_4$	Wald	-2LL
I. Qualidade do ensino	-0,14	0,35*	-	-	-	-	-	-	180,92
II. Qualidade do ensino + reprovação	-0,17	0,51*	0,76	3,92**	-	-	-	-	177,01
III. Qualidade do ensino + reprovação + já negociou débitos	-0,07	0,07*	0,48	1,37*	2,04	21,15***	-	-	152,19
IV. Qualidade do ensino + reprovação + já negociou débitos + possui cartão de crédito	-0,09	0,11*	0,63	2,13*	2,18	22,37***	-1,26	8,17***	143,41
V. Já negociou débitos + possui cartão de crédito	2,22	23,86***	-1,18	7,56***	-	-	-	-	145,59
VI. Já negociou débitos + possui cartão de crédito + reprovação	2,19	22,74***	-1,26	8,14***	0,61	2,05**	-	-	143,52
VI. Já negociou débitos + possui cartão de crédito + qualidade do ensino	2,21	23,58***	-1,18	7,57***	-0,04	0,03*	-	-	143,56

Fonte: dos autores

(\*) Estatisticamente não significativo

(\*\*) Estatisticamente significativo com 5%

(\*\*\*) Estatisticamente significativo com 1%

não será afetada e sua interpretação é semelhante ao teste *t-student*.

Analisando as variáveis explicativas a partir do Teste de Colinearidade, foram obtidos seis modelos de estimação apresentados na Tabela 3:

Como se pode observar na Tabela 3, o modelo número VI foi o que deu origem à função matemática correspondente ao modelo proposto, da qual fazem parte as variáveis X23 (reprovação), X30 (já negociou débitos) e X57 (possui cartão de crédito).

Comparando-se os resultados dos testes de avaliação do ajuste do modelo entre as propostas obtidas, verifica-se que o modelo VI apresentou *log da verossimilhança* (-2LL) no valor de 143,52 e todas as variáveis independentes com os coeficientes do modelo, apresentando significância a um nível de 0,01 e 0,05 a partir do Teste Wald.

## 4 ANÁLISE DOS RESULTADOS

### 4.1 Capacidade de previsão do modelo

Em modelos de regressão logística, a variável dependente é dicotômica ou binária; isto é,

pode assumir valor zero ou um. Neste estudo, foi atribuído valor zero para indicar a probabilidade de o aluno tornar-se adimplente e valor um para a probabilidade de o aluno se tornar inadimplente.

O ponto de corte utilizado no modelo foi de 0,25. Segundo Araújo e Carmona (2009, p. 17), “o ponto de corte 0,5 é o valor padronizado para a técnica de regressão logística”. Entretanto, esta determinação (0,25) foi proposta em razão de sua eficiência em produzir melhores resultados discriminatórios; ou seja, o ponto de corte foi sugerido em virtude de ser aquele que minimizava os erros de classificação. Portanto, os alunos com resultados superiores a 0,25 são classificados como inadimplentes e os alunos com resultados inferiores a 0,25 são classificados como adimplentes.

A capacidade de previsão do modelo foi analisada pela construção de uma matriz de classificação, que mostra os alunos distribuídos corretamente como adimplentes e inadimplentes. Essa matriz encontra-se na Tabela 4, a seguir.

Conforme demonstrado pela matriz de classificação na Tabela 4, o nível de acerto do modelo desenvolvido foi de 82%, tendo sido classificados corretamente 169 dos 206 alunos que participaram da amostra. Do grupo dos adim-

**Tabela 4** – Matriz de classificação do modelo de risco de crédito.

Observado	Estimado		Total	Classificações corretas
	Adimplentes	Inadimplentes		
Adimplentes	144	30	174	82,8%
Inadimplentes	7	25	32	78,1%
Total	151	55	206	82,0%

Fonte: Elaborado pelos autores, baseado em Brito e Assaf Neto, 2008, p. 25.

plentes, foram classificados corretamente 144 alunos de um total de 174 e do grupo dos inadimplentes foram classificados corretamente 25 alunos de um total de 32; lembrando que a amostra de trabalho tinha uma representatividade de 85% e 15% de adimplentes e inadimplentes, respectivamente, de acordo com a realidade encontrada na instituição objeto de estudo.

O resultado obtido neste estudo quando comparado a resultados obtidos em outros estudos de mesmo segmento, análise de risco de crédito, pode ser considerado satisfatório; ou seja, auxilia no processo de gestão da concessão de crédito da instituição.

Como exemplo de outros estudos e seus resultados, tem-se Minussi (2001, p. 99) com poder de previsão de 97,8%, Brito e Assaf Neto (2008, p. 25) com 90% e Silva e Freitas (2005, p. 418) com 89%. A menor eficiência deste estudo no acerto das previsões em relação aos citados tem como causa o tamanho de amostra utilizada, uma vez que no uso da técnica de regressão logística quanto maior for a amostra utilizada, melhor será o poder de previsão do modelo.

#### 4.2 Validação do modelo

Para a validação do modelo de regressão logística estipulado neste estudo, foram criadas duas amostras: a primeira chamou-se de amostra de análise, da qual se obteve um resultado de 82% de classificação correta dos alunos, e a segunda foi chamada de amostra de validação. A amostra de validação foi obtida através da aplicação de um

novo questionário aos alunos, contendo apenas as três variáveis necessárias para a estimação do modelo (reprovação, já negociou débitos e possui cartão de crédito).

O processo de seleção dos alunos que fizeram parte da amostra de validação foi realizado da seguinte forma:

- aplicação do questionário em diversas turmas de forma aleatória. Nesta etapa, foram obtidos 119 respondentes;
- identificação dos alunos adimplentes (97 alunos) e inadimplentes (22 alunos) entre os 119 respondentes;
- verificação da condição de inadimplência estabelecida neste estudo (atrasos iguais ou superiores a 45 dias) para os 22 alunos inadimplentes da amostra de validação. Restaram desta análise apenas 13 alunos nesta situação;
- divisão dos grupos de estudo de validação (adimplentes e inadimplentes), na mesma proporção obtida na amostra de análise; ou seja, 85% e 15%, representados por, respectivamente, 74 e 13 alunos.

Assim, a amostra de validação foi representada por outros 87 alunos da própria instituição, diferentes dos alunos que faziam parte da amostra de análise, dos quais 74 alunos possuíam a condição de adimplência e 13 alunos mostravam a condição de inadimplência. A definição dos grupos foi estipulada conforme a amostra de análise; ou seja, manteve-se a relação estabelecida: 85% de adimplentes e 15% de inadimplentes.

**Tabela 5** – Matriz de classificação – Validação do modelo.

Observado	Estimado		Total	Classificações corretas
	Adimplentes	Inadimplentes		
Adimplentes	54	20	74	73,0%
Inadimplentes	10	3	13	23,1%
Total	64	23	87	65,5%

Fonte: Elaborado pelos autores, baseado em Brito e Assaf Neto, 2008, p. 27.

A Tabela 5 apresenta os resultados obtidos a partir da amostra de validação.

Na Tabela 5, verifica-se que o percentual de acerto acumulado foi de 65,5%, em que foram classificados corretamente 57 alunos da amostra de validação. Entre os alunos adimplentes foram corretamente classificados 54 alunos, correspondendo a 73% da amostra total, enquanto que do grupo dos alunos inadimplentes, foram corretamente classificados 3 alunos, correspondendo a 23,1% da amostra total.

Comparado ao estudo de Brito e Assaf Neto (2008, p. 27), cuja validação apresentou resultado de 88,3%, nota-se uma queda no poder preditivo do modelo, isto se explica em razão do tamanho da amostra de validação ser muito reduzida, conforme já mencionado.

## 5 CONSIDERAÇÕES FINAIS

O objetivo deste estudo foi o de construir um modelo econométrico para estimar a probabilidade do risco de inadimplência em uma Instituição Privada de Ensino Superior. Obteve-se uma amostra de 206 alunos, sendo 174 alunos adimplentes e 32 alunos inadimplentes, em uma proporção de 85% e 15%, respectivamente. A técnica estatística aplicada na estimativa do modelo foi a Regressão Logística.

Os resultados obtidos por meio da pesquisa indicaram que o modelo previu com eficácia 82% da amostra total de análise, o que pode ser considerado um resultado satisfatório em termos de poder de estimativa da probabilidade de

risco de inadimplência, visto que estudos anteriores apresentaram resultados entre 80% e 98%, porém com amostras bastante maiores (MINUSSI, 2001, p. 99; BRITO; ASSAF NETO, 2008, p. 25; SILVA; FREITAS, 2005, p. 418). Estabelecendo-se uma leitura mais simplista, é possível dizer que, se o modelo estimou corretamente 82% dos casos, a diferença; ou seja, 18% dos casos corresponderiam à probabilidade do risco de inadimplência da instituição, o que, conforme descrito acima, é o objetivo geral deste estudo.

Além da hipótese de estimação do número de acertos previsto pelo modelo, em uma análise baseada a partir de modelos *logit*, o sinal estabelecido pelos coeficientes de regressão tem papel fundamental na interpretação dos resultados. Com relação a esta pesquisa, duas variáveis apresentaram sinais positivos e uma variável mostrou sinal negativo. As variáveis que fizeram parte do modelo e revelaram sinal positivo para os seus coeficientes foram “Reprovação” e “Já negociou débitos” e a variável com sinal negativo foi “Possui cartão de crédito”.

Em termos práticos, a leitura do modelo para um aluno que apresentasse as três condições acima descritas seria a seguinte: o aluno que obteve reprovação durante seu período letivo na Universidade, que já negociou débitos e que não possui cartão de crédito, tende a se tornar um provável inadimplente. Observe que a condição para o sinal negativo do coeficiente da variável “Possui cartão de crédito” indica a ausência da posse do cartão de crédito.

A análise dos sinais dos regressores apresenta uma lógica esperada. O aluno com reprova-

ção possui custo maior para finalizar seus estudos, tendendo a ter também maiores dificuldades de pagamento. O aluno que já realizou composição de débitos anteriores apresenta perfil de dificuldade financeira. E, por último, o fato do aluno não possuir cartão de crédito indica, teoricamente, instabilidade na renda, baixo poder aquisitivo e, talvez, que tenha um histórico de crédito ruim.

Essa situação permite avaliar o bom o desempenho do modelo elaborado neste estudo, uma vez que Gujarati (2006, p. 488) entende que em modelos de regressão binários, a qualidade do ajuste tem importância secundária. O mais importante são os sinais esperados para os coeficientes de regressão e a sua significância estatística e/ou prática.

Por fim, embora este estudo não seja considerado um estudo definitivo em termos de probabilidade de estimativa do risco de inadimplência em UP, seus resultados e considerações estavam de acordo com a realidade encontrada, o que justifica a sua aplicabilidade. Porém, recomenda-se a observância de dois pontos importantes: o instrumento de coleta utilizado, embora apresente pontos positivos relevantes e dinâmicos ao processo de pesquisa, deve ser muito bem validado e seus vieses considerados no momento da coleta de dados, e a utilização do modelo estatístico não pode ser aplicada isoladamente para medir a probabilidade de risco de inadimplência.

Análises a respeito do comportamento do aluno (passado e presente) devem complementar a técnica estatística. Exemplo disto é um aluno ter incorrido em débitos no passado, devido a fatores como desemprego ou doença, situação transitória e já superada. A aplicação pura e simples do modelo, sem análise qualitativa provavelmente o classificaria como possível inadimplente.

Como sugestão para futuras pesquisas: a aplicação do modelo de estimação em outras UPs da mesma região ou diferente desta, comparando se os resultados obtidos tendem a similaridade ou a disparidade; além de obtenção de amostras de análise subdivididas por curso ou por turno (manhã, tarde ou noite), examinando a hipótese da existência de classes sociais diferentes para cada curso e avaliando seu reflexo na análise do risco de inadimplência.

## REFERÊNCIAS

ARAÚJO, E. A.; CARMONA, C. U. de M. Construção de Modelos Credit Scoring com Análise Discriminante e Regressão Logística para a Gestão do Risco de Inadimplência de uma Instituição de Microcrédito. **REAd: Revista Eletrônica de Administração**. Porto Alegre. ed. 62, v. 15, n. 1, jan./abr. 2009. Disponível em: <[http://www.read.ea.ufrgs.br/edicoes/pdf/artigo\\_579.pdf](http://www.read.ea.ufrgs.br/edicoes/pdf/artigo_579.pdf)>. Acesso em: 14 fev. 2011

BANCO CENTRAL DO BRASIL. Boletim do Banco Central – Relatório Anual 2006. Brasília. Disponível em: <<http://www.bcb.gov.br/?relouvid2006>>. Acesso em: 20 nov. 2006.

BRITO, G. A. S.; ASSAF NETO, A. Modelo de classificação de risco de crédito de empresas. **Revista Contabilidade & Finanças USP**, São Paulo, v.19, n. 46, p. 18–29, jan./abr. 2008.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. **Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, 2007.

FREITAS, H. et al. Pesquisa via Internet: características, processo e interface. **Revista Eletrônica GIANTI**, Porto Alegre, 2004, 11p. Disponível em: <[http://www.ea.ufrgs.br/professores/hfreitas/files/artigos/2010/2010\\_gianti\\_hf\\_rsc\\_pesquisa\\_pontual\\_ou\\_escuta\\_permanente.pdf](http://www.ea.ufrgs.br/professores/hfreitas/files/artigos/2010/2010_gianti_hf_rsc_pesquisa_pontual_ou_escuta_permanente.pdf)>. Acesso em: 17 mar. 2011

GONÇALVES, E. B. **Análise de risco de crédito com uso de modelos de regressão logística, redes neurais e algoritmos genéticos**. 2005. 105 f. Dissertação (Mestrado em Administração de empresas) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. São Paulo, 2005.

GUIMARAES, I. A.; CHAVES NETO, A. Reconhecimento de padrões: metodologias estatísticas em crédito ao consumidor. **RAE electron**. [online], v. 1, n. 2, p. 02-14, jul./dez. 2002. Dis-

ponível em: < <http://www.scielo.br/pdf/raeel/v1n2/v1n2a06.pdf>>. Acesso em: 31 jan. 2011.

GUJARATI, D. N.. **Econometria básica**. 4. ed. Rio de Janeiro: Elsevier, 2006.

HAIR JUNIOR, J. F. et al. **Análise multivariada de dados**. 5. ed. Porto Alegre: Bookman, 2005.

JOLY, M. C. R.; SILVEIRA, M. A. Avaliação preliminar do questionário de informática educacional (QIE) em formato eletrônico. **Psicologia em Estudo**, Maringá, v. 8, n. 1, p. 85-92, jan./jun. 2003. Disponível em: < <http://dx.doi.org/10.1590/S1413-73722003000100011>>. Acesso em: 15 jan. 2011

LOCH, J. M.; REIS, D. R. A expansão do ensino superior: um comparativo do crescimento no Brasil, no Paraná e na região metropolitana de Curitiba. In: **WORLD CONGRESS ON ENGINEERING AND TECHNOLOGY EDUCATION – WCETE'2004**, Santos, 2004. v. 1 . Anais... Santos: UniSantos, 2004. p. 58.

MINUSSI, J. A.. **Modelo preditivo de solvência utilizando regressão logística**. 2001. 150 f. Dissertação (Mestrado em Administração de Empresas) – PUC-RIO/UNISINOS, São Leopoldo, 2001.

QUEIROZ, R. S. B. de. A importância dos modelos de *Credit Scoring* na concessão de crédito ao consumidor no varejo. In: **SEMINÁRIO DE ADMINISTRAÇÃO DA USP**, 9, 2006, São Paulo. Anais... São Paulo: PPGA-USP, 2006.

SANDRONI, P. **Dicionário de economia do século XXI**. 2. ed. Rio de Janeiro: Record, 2006.

SECURATO, J. R.. **Crédito – análise e avaliação do risco: pessoas físicas e jurídicas**. São Paulo: Saint Paul Institute of Finance, 2002.

SIEGEL, S. **Estatística não paramétrica para as Ciências do Comportamento**. São Paulo: McGraw-Hill do Brasil, 1975.

SILVA, W. V.; FREITAS, D. D. de. Estimação do modelo de risco de inadimplência dos discentes regularmente matriculados na PUCPR usando a técnica de Análise Discriminante. In: **SIMPÓSIO DE EXCELÊNCIA EM GESTÃO E TECNOLOGIA – SEGeT**, 2, 2005. Anais... Resende: AEDB, 2005.

THOMAS, L. A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. **International Journal of Forecasting**, Amsterdam, v. 16, n. 2, p. 149-172, Apr./Jun., 2000.

## APÊNDICE

## Apêndice A – Questionário de coleta de dados para análise de risco de crédito pessoa física

## Dados pessoais do aluno:

1. Sexo.
  - a)  Masculino
  - b)  Feminino
2. Faixa etária.
  - a)  Abaixo de 20 anos
  - b)  De 20 anos até 25 anos
  - c)  Acima de 25 anos
3. Qual a sua naturalidade?
  - a)  Pelotas
  - b)  Outras localidades
4. Reside em casa própria?
  - a)  Sim
  - b)  Não
5. Com quem reside?
  - a)  Família
  - b)  Parentes
  - c)  Outros
6. Estado civil?
  - a)  Solteiro(a)
  - b)  Casado(a) ou morando com  
companheiro(a)
  - c)  Outro
7. Possui filhos ou dependentes?
  - a)  Sim
  - b)  Não
8. Em caso afirmativo à pergunta anterior, quantos?
  - a)  De zero a um
  - b)  Mais do que 1 até 3
  - c)  Mais do que 3

## Informações acadêmicas e financeiras do aluno:

9. Quantas disciplinas você já cursou (responda desconsiderando as disciplinas que está cursando)?
  - a)  De 3 a 10
  - b)  Mais de 10 até 20
  - c)  Mais do que 20
10. Quantas disciplinas faltam cursar? (Incluir as disciplinas que está cursando atualmente)
  - a)  De 3 a 10
  - b)  Mais de 10 até 20
  - c)  Mais do que 20
11. Possui alguma reprovação?
  - a)  Sim
  - b)  Não
12. Em caso afirmativo à pergunta anterior, em quantas disciplinas?
  - a)  Até duas
  - b)  Mais do que duas
13. Possui alguma mensalidade em atraso?
  - a)  Sim
  - b)  Não
14. Em caso afirmativo à pergunta anterior, qual o valor do débito em atraso?
  - a)  Até R\$ 1.000,00
  - b)  Acima de R\$ 1.000,00 até R\$ 5.000,00
  - c)  Acima de R\$ 5.000,00
15. Você já negociou débitos com a central de negociações da universidade?
  - a)  Sim
  - b)  Não
16. Em caso afirmativo à pergunta anterior, quantas vezes negociou débitos?
  - a)  Até duas vezes
  - b)  Mais do que duas vezes

**Informações profissionais do aluno:**

17. Possui alguma atividade remunerada (estágio, bolsa de estudos, atividade com vínculo empregatício, etc.)?

- a)  Sim
- b)  Não

18. Qual o valor que recebe por essa atividade?

- a)  Até R\$ 415,00
- b)  Acima de R\$ 415,00 até R\$ 4.150,00
- c)  Acima de R\$ 4.150,00

**Dados referentes ao grupo familiar:**

19. Qual o nível de escolaridade do seu pai?

- a)  Sem escolaridade
- b)  Ensino Fundamental ou Médio (completo ou incompleto)
- c)  Ensino Superior ou Pós-Graduação (completo ou incompleto)

20. Qual o nível de escolaridade da sua mãe?

- a)  Sem escolaridade
- b)  Ensino Fundamental ou Médio (completo ou incompleto)
- c)  Ensino Superior ou Pós-Graduação (completo e incompleto)

21. Existem pessoas que moram com você e que dependam da renda da família?

- a)  Sim
- b)  Não

22. Em caso afirmativo à pergunta anterior, quantas pessoas dependem dessa renda?

- a)  Até duas
- b)  Mais do que duas

23. Qual a renda da família (soma dos salários de todas as pessoas que possuam atividade remunerada)?

- a)  Até R\$ 415,00
- b)  Acima de R\$ 415,00 até R\$ 4.150,00
- c)  Acima de R\$ 4.150,00

24. Tem familiares cursando o ensino superior?

- a)  Sim
- b)  Não

25. Em caso afirmativo à pergunta anterior, quantos estudam em instituição privada?

- a)  Até dois
- b)  Mais do que dois

**Dados sobre as responsabilidades financeiras do aluno:**

26. Possui algum tipo de financiamento, como imóvel, veículo ou empréstimo bancário?

- a)  Sim
- b)  Não

27. Em caso afirmativo à pergunta anterior, qual o valor desse comprometimento financeiro? (Considere a soma deles)

- a)  Com prestação no valor até R\$ 415,00
- b)  Com prestação no valor acima de R\$ 415,00 até R\$ 4.150,00
- c)  Com prestação acima de R\$ 4.150,00

28. Possui conta bancária?

- a)  Sim
- b)  Não

29. Possui cartão de crédito?

- a)  Sim
- b)  Não

30. Possui seguro de vida?

- a)  Sim
- b)  Não

**Dados sobre a IES do aluno:**

31. Qual nota você daria para o nível de qualidade de ensino de sua instituição, considerando 1 a menor nota e 5 a maior nota?

- a)  Nota 1
- b)  Nota 2
- c)  Nota 3
- d)  Nota 4
- e)  Nota 5