

ANÁLISE DE INADIMPLÊNCIA A PARTIR DO MODELO DE *CREDIT SCORE* APLICADO A PESSOAS FÍSICAS DE BAIXA RENDA.

RESUMO: Os modelos de *Credit Scoring* são modelos quantitativos empregados comumente por instituições financeiras na mensuração e previsão de risco de crédito. Nesse contexto, o presente trabalho analisa o nível de inadimplência dos consumidores pessoa física de baixa renda de Curitiba e região metropolitana e a partir da resolução 2.682/99 do Banco Central do Brasil estima um modelo de *Credit Score* para inadimplência. Na análise foi utilizada a Regressão Logística com o auxílio dos *softwares SPSS e Xlstat*. A partir dos resultados obtidos por meio da análise quantitativa observou-se que dos 2.369 respondentes que formam a amostra do estudo, 335 são considerados maus pagadores, por possuírem três parcelas ou mais em atraso o que corresponde ao prazo acima de 90 dias. No que se refere à construção do modelo de previsão de inadimplência, utilizou-se a técnica de regressão logística binária. Considerando que o percentual de inadimplentes foi igual a 14,14%, (335 respondentes) selecionou-se uma amostra aleatória de 200 indivíduos para a amostra de validação, com o objetivo de verificar a aderência do modelo estimado aos dados observados. Esta amostra possibilita comprovar se o modelo está predizendo os dados adequadamente. De uma forma geral todos os resultados obtidos na pesquisa foram satisfatórios.

PALAVRAS-CHAVE: Análise de Crédito; Risco de Crédito; Endividamento.

ABSTRACT: Credit Scoring models are quantitative commonly employed by financial institutions in the measurement and prediction of credit risk. This paper analyzes the individual consumer default level of low-income in Curitiba and metropolitan region from resolution 2.682/99 of Central Bank of Brazil that estimates a model for default Credit Score. The analysis was used logistic regression with the help of SPSS and Xlstat. From the results obtained by quantitative analysis It was noticed that the 2,369 respondents who form the study sample, 335 are considered bad payers, by having three payments or more past due which corresponds to the period over 90 days. In what refers to the construction of the default prediction model, was used the binary logistic regression technique. Considering the percentage of defaulters was equal to 14.14% (335 respondents), It was selected a random sample of 200 individuals for the validation sample, with the objective to verify the compliance of the estimated model to the observed data. This sample provides evidence that the model is predicting the data adequately. In general, all the results obtained in the study were satisfactory.

KEY-WORDS: Credit Analysis; Credit Risk; Indebtedness.

1 INTRODUÇÃO

No Brasil, antes de 1994 com os famosos “pacotes econômicos”, as operações do mercado de crédito se alteravam e tinham sérias implicações sociais (SICSÚ, 2010). A partir de 1994, com a implantação do Plano Real e a estabilização da moeda, a concessão de crédito financeiro passou a ser um negócio cada vez mais rentável para as instituições bancárias e socialmente mais viável aos usuários, uma vez que estas já não contavam com os ganhos oriundos da desvalorização da moeda (CAMARGOS et al., 2012).

De acordo com Ribeiro, Zani e Zanini (2009), as facilidades na concessão de crédito a pessoas físicas têm aumentado no decorrer dos anos. Variáveis como redução das taxas de juros, aumento dos prazos de pagamentos, empréstimos consignados à folha de pagamento e o crescimento da renda possibilitaram à população em geral acesso a aquisição de diversos bens

e serviços antes inatingíveis para determinadas classes sociais.

Embora o cenário brasileiro, apresente uma nova forma de relacionamento com o crédito de varejo, a concessão, não poderia ser feita a todos os demandantes, razão pela qual foi necessário o aprimoramento dos procedimentos de avaliação, para a tomada de decisão adequada sobre emprestar ou não capital ao proponente (ROCHA, et al, 2010). Com isso, e diante da morosidade e da ausência de critérios objetivos, as instituições financeiras de crédito começaram a adotar modelos de análise para concessão de crédito, com o objetivo de acelerar o processo de aprovação e minimizar o viés causado pela subjetividade do antigo processo, promovendo, assim, maior objetividade e agilidade de decisão, menores custos, melhor predição e conseqüentemente menor risco (NOBRE, et al, 2017).

Ribeiro, Zani e Zanini (2009) destacam que, na maioria dos mercados, existe uma correlação positiva entre o crescimento da política de concessão de crédito e o aumento dos níveis de inadimplência. Neste contexto, a cada dia se tornaram mais necessárias ferramentas de análise de risco de crédito, mais robustas, seguras e confiáveis, no sentido de evitar ou reduzir os níveis de inadimplência e que traduza as informações sobre os clientes tomadores de crédito.

Nesse contexto, o presente artigo tem como objetivo central, analisar o nível de inadimplência de pessoas físicas de baixa renda da cidade de Curitiba e região metropolitana, a partir da resolução 2.682/99 do Banco Central do Brasil por meio do modelo de *credit score*. Para tanto, apresenta a seguir a abordagem teórica, metodologia, análise dos dados e considerações finais.

2 ABORDAGEM TEÓRICO-EMPÍRICA: MODELAGEM DE CRÉDITO

Historicamente, no Brasil, a concessão de crédito sempre foi lenta e escassa, devido a políticas mal concebidas e ao processo inflacionário do passado. Entretanto, em consequência da maior estabilidade da economia brasileira nos últimos anos, após a implantação do Plano Real, as empresas têm percebido o crédito como um gerador de riquezas e de novos negócios (BUENO, 2003).

De acordo com Santos (2009), o crédito inclui as noções fundamentais de confiança, expressa na promessa de pagamento, e tempo, que se refere ao período fixado entre a aquisição e a liquidação da dívida. Para o autor a confiança é um elemento necessário, porém, não é suficiente para uma decisão de crédito.

Nesse sentido, no mercado de crédito, saber escolher os clientes pode ser a chave para o sucesso do negócio em qualquer ambiente. De acordo com Louzada et al. (2009), a classificação de cliente é de vital importância para a determinação de concessões.

Camargos et al. (2012), sugerem que o enquadramento dos tomadores ou das operações de crédito nos níveis de risco deve ocorrer fundamentado em aspectos quantitativos e qualitativos, pois os critérios precisam ser predeterminados com base técnica, a fim de evitarem avaliações divergentes da política de crédito.

Nas últimas décadas o progresso tecnológico viabilizou técnicas estatísticas avançadas para análise de crédito, de tal maneira que atualmente é o mais importante suporte para tomada de decisões ante ao enorme volume de propostas de crédito, advindas especialmente de pessoas físicas (SANTOS, 2009).

Silva (2006) apresenta diversas ferramentas possuidoras de um suporte estatístico significativo que são usadas na avaliação do risco de crédito, a saber: os sistemas especialistas, *credit scoring*, *behaviour scoring*, *rating*, redes neurais, algoritmos genéticos, dentre outras.

Instituições financeiras de muitos países, entre eles o Brasil, estão intensificando e aperfeiçoando metodologias de estudos sobre práticas que auxiliam no controle dos riscos, tendo como ferramentas poderosas para apoio à decisão de crédito, os métodos estatísticos e as abordagens de inteligência artificial (THOMAS, 2000).

Esses métodos são usados na elaboração dos modelos de *Credit Scoring* que consistem em uma das principais ferramentas de suporte à concessão de crédito. O desenvolvimento desses modelos, segundo Louzada et al (2009), baseia-se na construção de um procedimento formal para determinar quais características do cliente estão relacionadas, significativamente, com o seu risco de crédito e qual a intensidade e direção desse relacionamento. O objetivo básico é geração de uma pontuação, pela qual os clientes possam ser classificados conforme a sua chance de inadimplência.

De acordo com Sicsú (2010), os modelos de *credit scoring* podem ser aplicados às solicitantes de crédito com os quais o credor não teve nenhuma experiência anterior, enquanto que o *behavioural scoring* são modelos desenvolvidos para clientes ou ex-clientes, ou seja, solicitantes com um histórico de crédito em operações com o credor. A diferença entre esses modelos reside nas variáveis utilizadas para estimar o risco do crédito.

Dentre as técnicas estatísticas utilizadas nos modelos de *credit scoring*, pode-se citar: a análise discriminante e a regressão logística, sendo que a regressão logística será utilizada nessa pesquisa (ROCHA, et al 2010).

Desde 1960 que o *Credit Scoring* tem revolucionado profundamente os processos de decisão de crédito. O *Credit Scoring* representa uma importante ferramenta de estimação e redução de risco de crédito. O qual passou a ser largamente utilizado, tanto no Brasil quanto no exterior, em análise de crédito direto tanto de pessoa física quanto pessoa jurídica, sendo inicialmente utilizado por instituições bancárias e seguradoras e atualmente é utilizado por vários segmentos de mercado dentre eles o mercado de varejo (THOMAS, 2000).

De acordo com o autor citado anteriormente, modelos de *Credit Scoring* são sistemas que atribuem pontuações às variáveis de decisão de crédito de um proponente, mediante a aplicação de técnicas estatísticas. Esses modelos visam a segregação de características que permitam distinguir os bons dos maus créditos. O autor, afirma que partir de uma equação gerada por variáveis referentes ao proponente de crédito e/ou à operação de crédito, os sistemas de *Credit Scoring* geram uma pontuação que representa o risco de perda. O escore que resulta da equação de *Credit Scoring* pode ser interpretado como probabilidade de inadimplência ao se comparar a pontuação de um crédito qualquer com determinada pontuação estabelecida como ponto de corte ou pontuação mínima aceitável. Para Thomas (2000), a pontuação de crédito é um instrumento estatístico desenvolvido para que o analista avalie a probabilidade de que determinado cliente venha a tornar-se inadimplente no futuro.

De acordo com Sicsú (2010), os escores de crédito tem a finalidade única de quantificar o risco de crédito e que a forma em que essa informação é utilizada para tomar a decisão de conceder ou não o crédito, para estruturar a operação definindo taxas, garantias, prazos, etc., é atribuição dos gestores de crédito. Portanto, para operacionalizar um processo de concessão e gestão de crédito faz-se necessário, além do modelo de cálculo dos escores, uma política de crédito bem definida, um sistema de informações gerenciais com dados do cliente, operação, formas de pagamento, políticas de cobrança, etc (SICSÚ, 2010).

3 ABORDAGEM TEÓRICO-EMPÍRICA: INADIMPLÊNCIA E MERCADO BRASILEIRO

As classes sociais permitem uma ordenação dos consumidores e de suas famílias com o compartilhamento de determinados valores, crenças, interesses e estilos de vida que

refletem nos desejos de compra. Assim, o comportamento do consumidor pode ser influenciado por interesses compartilhados por uma determinada classe social e especialmente se os consumidores almejam a ascensão social para diferenciar-se de seus grupos com objetivo da elevação do status social (SILVA, 2012).

Segundo Casado (2000), sob a concepção atual da sociedade em que se vive, infelizmente, o homem é o resultado do seu patrimônio. Sua casa, suas vestimentas, seus acessórios, seu carro e todos os demais bens de consumo. Tudo isto é o que traduz a existência humana. A pessoa ocupa determinada classe social dependendo de suas posses. A sociedade de consumo padroniza costumes e cria necessidades (ANDRICH, et al, 2014).

Como o Brasil tem disparidades regionais e de renda, a Associação Brasileira de Empresas de Pesquisa (ABEP) classifica o mercado de acordo com os hábitos de consumo e a escolaridade dos chefes de família. Existem seis classes sociais: A; B1; B2; C1; C2 e DE.

Conforme a ABEP (2013), em função do tamanho reduzido da Classe A1 a renda média deixa de ser divulgada. Sendo assim, a estimativa de renda média é feita para o conjunto da classe A. Bem como, devido o tamanho reduzido da classe E, que inviabiliza a leitura de resultados obtidos por meio de amostras probabilísticas ou por cotas, as classes D e E foram unidas para a estimativa e construção de amostras.

Para a classificação dos domicílios são utilizados como fonte de dados informações sobre a posse de bens duráveis (TV, máquina de lavar, geladeira, etc.), o nível de instrução, o número de banheiros na casa e se a família conta ou não com empregada mensalista.

O objeto da presente pesquisa, conforme descrito na introdução, está vinculado a pessoas denominadas de “baixa renda”, esse termo, embora comum na classificação econômica de povos, apresenta grandes diferenças em todo mundo, no que tange a renda da população (MOREIRA, 2006). Conforme classificação da (ABEP, 2013), pessoas das classes sociais C1, C2 e DE são consideradas como população de baixa renda no Brasil. Portanto, essa parcela da população representa um importante segmento para a economia do país, representando uma interessante oportunidade de mercado, principalmente para empresas que estão buscando novas alternativas de crescimento. Vale ressaltar que para definir baixa renda não depende apenas de quanto as pessoas ganham, mas também do estilo de vida que as mesmas levam.

Os consumidores de baixa renda têm sido, tradicionalmente os alvos das redes de varejo e indústrias de diferentes setores, os mesmos estão lutando para conquistar um pedaço da renda da participação de consumo da população de baixa renda e procuram atender, cada vez melhor, o cliente desse segmento (MOREIRA, 2006).

Para ajudar a ampliar a possibilidade de compra tanto de necessidades quanto de desejos, as empresas de varejo utilizam o crédito. De acordo com Brusky e Fortuna (2002), para a população de baixa renda, comprar a crédito significa ganhar um prazo para poder reunir as condições necessárias para efetuar os pagamentos, de forma que o que está sendo negociado é o tempo. Uma preocupação é manter o “nome limpo”, não constando dos registros negativos das instituições de proteção ao crédito, para assim continuar comprando e tendo acesso ao crédito formal.

4 PROCEDIMENTOS METODOLÓGICOS

A presente pesquisa caracteriza-se como descritiva e exploratória, com dados primários analisados por meio de técnicas estatísticas univariadas e multivariadas valendo-se do pacote estatístico SPSS e *Xlstat*. No que se refere a natureza da pesquisa, é de natureza aplicada. Em relação aos objetivos, esta pesquisa caracteriza-se ainda por ser relacional, já que visa determinar a influência das variáveis independentes explicativas sobre a variável

dependente de saída, que se trata de variável dicotômica classificando a amostra analisada em consumidores pessoa física adimplentes ou inadimplentes. Quanto aos procedimentos técnicos, para investigação, serão utilizadas as pesquisas do tipo, bibliográfica e *survey*.

A população do estudo são os consumidores pessoas físicas residentes na cidade de Curitiba e região metropolitana, a qual corresponde a 3,4 milhões (três milhões e quatrocentos mil habitantes) conforme o último Censo do IBGE realizado em 2010, sendo que desses, aproximadamente 70% pertencem as classes sociais C1, C2 e DE as quais são classificadas como baixa renda. Sendo que a amostra coletada para esta pesquisa é composta por 2.369 pessoas físicas de baixa renda, escolhida a partir de indivíduos que possuem a idade mínima de 18 anos de idade, em função da sua maioria penal, residentes em diversos bairros da cidade de Curitiba onde se concentra a maior parte das unidades amostradas, bem como nos municípios da região metropolitana de Curitiba. Sendo que, 1332 respondentes residem em Curitiba o que corresponde a 56% da amostra e 1037 residem nos municípios da região metropolitana correspondendo a 44% da amostra. O método de amostragem utilizado nessa pesquisa é do tipo não probabilístico por acessibilidade.

Os dados foram obtidos por meio de questionário sócio econômico aplicado para consumidores pessoa física de Curitiba e região metropolitana, durante os meses de agosto a outubro de 2013, sendo aplicados para as pessoas físicas que estavam dispostas a responder. Tais dados foram submetidos ao tratamento de Regressão Logística, gerando resultantes passíveis do cumprimento do objetivo central da pesquisa.

5 APRESENTAÇÃO E ANÁLISE DOS DADOS

5.1 DEFINIÇÃO DOS GRUPOS DE MAUS E BONS PAGADORES

Em 22 de dezembro de 1999, o Conselho Monetário Nacional por intermédio do Banco Central do Brasil aprova a Resolução 2.682 que surge para regular a análise de risco efetuada pelas instituições financeiras. A Resolução 2.682/99 define que as operações de crédito das instituições financeiras devem ser analisadas conforme dois critérios: (a) por ordem crescente de risco, onde “AA” representa risco zero e “H” o maior risco e (b) por dias vencidos conforme pode ser visto no Quadro 1.

Quadro 1 – Critério de Classificação da Resolução 2.682/99

Categoria	AA	A	B	C	D	E	F	G	H
Atraso (dias)	0	0-14	15-30	31-60	61-90	91-120	121-150	151-180	> 180
Percentual	0%	0,5%	1%	3%	10%	30%	50%	70%	100%

Conforme descrito no Quadro 1, os dias em atraso considerados, variam de “0” um cliente pontual, classificado com um *rating* de AA até o mínimo de 180 dias onde o *rating* recebido é de “H”. Vale destacar que os valores percentuais em todos os casos se referem à provisões e reservas de recursos para as prováveis perdas na realização do crédito, os quais são sobre o valor das operações classificadas conforme o nível de risco de cada categoria.

A partir dos resultados obtidos por meio dos questionários e com base na Resolução 2682/99, foi arbitrado, por conveniência que bom cliente seria aquele **com atraso até 90 dias**. Sendo que para o estudo foram considerados inadimplentes ou cliente mau pagador os respondentes que declararam possuir dívidas **com atraso superior a 90 dias**, inclusive os que possuem dívidas acima de 5 anos, sendo que a partir deste prazo as restrições são retiradas do CPF do indivíduo, porém, como ele ainda não liquidou a dívida ele continua inadimplente e o registro permanece nos sistemas de seu credor.

Para selecionar os bons dos maus clientes, seguiram-se dois passos: primeiro, verifica-se o período de atraso, seguido do agrupamento dos clientes entre “bons” e “maus” de acordo com o histórico de atrasos.

A Tabela 1 resume a situação dos respondentes em função do tempo de atraso de seus compromissos financeiros.

Tabela 01 - Classificação do Respondente em Bom ou Mau de Acordo com o Tempo de Atraso

Gênero	% que possui dívidas em atraso	Número de Parcelas em Atraso						
		Não possui dívidas	Nenhum a parcela	1	2	3	4 ou mais	Acima de 5 anos
Masc.	Contagem	504	316	63	36	32	68	43
	%	42,2%	51,9%	45,7%	38,7%	41,0%	41,2%	46,7%
Fem.	Contagem	690	293	75	57	46	97	49
	%	57,8%	48,1%	54,3%	61,3%	59,0%	58,8%	53,3%
Total	Contagem	1.194	609	138	93	78	165	92
	%	100%	100%	100%	100%	100%	100%	100%

Percebe-se que dos 2.369 respondentes, 1.194 não possuem dívidas, 609 possuem dívidas, porém não estão com nenhuma parcela atrasada e 566 respondentes declararam possuir dívidas em atraso, sendo que desses 231 foram considerados como bons clientes, pois no momento da coleta de dados declararam possuir uma ou duas parcelas em atraso o que corresponde ao prazo de até 90 dias enquanto que 335 são considerados maus pagadores, por possuírem três parcelas ou mais em atraso o que corresponde ao prazo acima de 90 dias.

5.2 EXCLUSÃO DAS VARIÁVEIS INDEPENDENTES

A inclusão ou exclusão das variáveis no modelo de risco de crédito (*Credit Score*) dependerá dos resultados estatísticos dispostos nos testes de hipóteses empregados. Nesse sentido, testes de hipóteses de associação do qui-quadrado, a função causa-raiz, os coeficientes de correlação não-paramétrica de *spearman* e o coeficiente de explicação ou de determinação (R^2) são apresentadas na pesquisa, visando definir quais as variáveis independentes discriminam o fato do respondente ser caracterizado como um bom ou mau pagador.

Nesse sentido, foram excluídas as variáveis demográficas que apresentam a simbologia (#) devido ao fato do teste de hipóteses de qui-quadrado (χ^2) não ter sido estimado, pois as questões apresentavam múltiplas respostas.

A Tabela 2 traz de forma sumarizada quais as variáveis independentes candidatas a fazerem parte do modelo a ser estimado.

Tabela 2 - Exclusão das Variáveis a Partir da Estatística Qui-Quadrado

Variável	χ^2	Valor-p	Inclusão/ Exclusão
Q1 - Gênero	0,724	0,395	Exclusão
Q2 - Idade	45,878 ^a	0,000	Inclusão
Q3 - Estado Civil	17,513 ^a	0,004	Inclusão
Q4 - Maior escolaridade completa	28,044	0,000	Inclusão

Q5 - Número de filhos	70,764 ^a	0,000	Exclusão
Q5.1 - Numero de filhos dependentes	72,474 ^a	0,000	Inclusão
Q6 - Idade dos filhos	#	#	Exclusão
Q7 - Situação profissional	13,122 ^a	0,069	Exclusão
Q8 - Tempo de trabalho no atual emprego	14,876 ^a	0,011	Inclusão
Q9 - Renda mensal individual	24,889 ^a	0,000	Inclusão
Q10 - Número de pessoas dependentes	56,746 ^a	0,000	Inclusão
Q11-Percentual da renda gasta com necessidades básicas	33,232 ^a	0,000	Inclusão
Q12 - Atividades que realiza nos dias de folga	#	#	Exclusão
Q13 - Local onde reside	5,790 ^a	0,215	Exclusão
Q14 - Tipo de moradia	44,101 ^a	0,000	Inclusão
Q15 - Tempo que reside no atual endereço	38,600 ^a	0,000	Inclusão
Q16 - Número de pessoas que residem no mesmo imóvel	8,616 ^a	0,125	Exclusão
Q17 - Tipos de seguro que possui	#	#	Exclusão
Q18 - Meios de transporte que utiliza	#	#	Exclusão
Q19 - Classificação econômica	1,040 ^a	0,959	Exclusão
Q20 - Grau de instrução do chefe de família	22,908 ^a	0,003	Inclusão
Q21 - Tipo de conta bancária	22,747 ^a	0,001	Inclusão
Q22 - Limite do cartão de crédito	41,052 ^a	0,000	Inclusão
Q23 - Ocasões em que utiliza o limite da conta corrente	114,232 ^a	0,000	Inclusão
Q24 - Formas de pagamento do cartão de crédito	95,664 ^a	0,000	Inclusão
Q25 - Valor da parcela de empréstimo e/ou financiamento	61,913 ^a	0,000	Inclusão
Q26 - Adquiriu ou não imóvel	1,188 ^a	0,276	Exclusão
Q26.1 - Tipo de bem imóvel adquirido nos últimos três anos	2,905 ^a	0,406	Exclusão
Q27 - Saldo a vencer do empréstimo e/ou financiamento	100,263 ^a	0,000	Inclusão
Q28 - Número de parcelas que restam para pagar do empréstimo e/ou financiamento	65,654 ^a	0,000	Inclusão
Q29 - Formas de pagamento quando realiza compras	36,996 ^a	0,000	Inclusão
Q30 - O que influenciou a realizar compras sem necessidade	19,429 ^a	0,001	Inclusão
Q31 - Número de parcelas de dívida em atraso	<i>STATUS</i>	<i>STATUS</i>	<i>STATUS</i>
Q32 - Motivos que levaram a possuir dívidas em atraso.	1050,763 ^a	0,000	Inclusão
Q33 - Número de vezes que renegociou dívidas em atraso	238,699 ^a	0,000	Inclusão
Q34 - Para quem o respondente emprestou o crédito	28,520 ^a	0,000	Inclusão
Q35 - De quem o respondente emprestou o crédito	48,479 ^a	0,000	Inclusão
Q36 - Ocasões em que o nome foi incluído no sistema de proteção ao crédito	630,078 ^a	0,000	Inclusão
Q37-Conhecimento das consequências quando o nome é incluído no SPC	5,361 ^a	0,069	Exclusão
Q38 - Como o respondente cuida da saúde	14,876 ^a	0,011	Inclusão

A partir da Tabela 2 é possível observar a exclusão, com base na estatística de associação do qui-quadrado, 13 variáveis independentes ou explicativas. Notadamente que outras métricas estatísticas são usadas visando avaliar posterior exclusão.

5.3 EXCLUSÃO DAS VARIÁVEIS A PARTIR DA FUNÇÃO CAUSA RAIZ

A partir das variáveis remanescentes da seção anterior, foi realizado o teste de hipóteses da Função Causa Raiz, onde avaliou por meio da estatística F-Fisher, observando a importância que cada variável independente em relação a variável dependente.

A Tabela 3 mostra as variáveis que apresentaram maior importância são as que: analisa se o respondente possui dívidas em atraso e qual foi o motivo do atraso (F-Fisher = 394,19), analisa se o nome já foi incluído no sistema de proteção ao crédito (SEPROC/SPC),

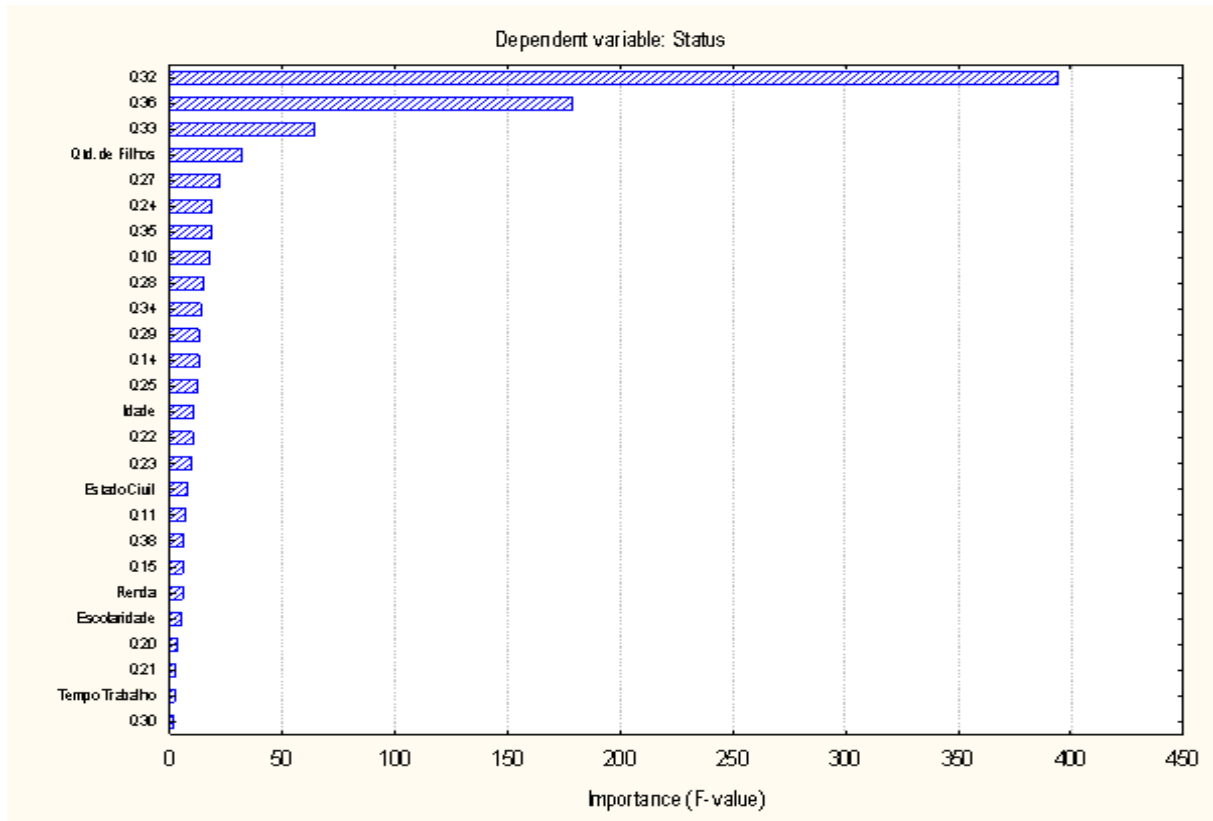
apresentou um valor igual a 178,92 (F-Fisher = 178,92), analisa se o respondente renegociou débitos (dívidas em atraso), a qual apresentou uma estatística F-Fisher = 64,52, analisa o número de filhos do respondente, a qual apresentou F-Fisher 31,91, analisa o valor total do empréstimo ou financiamento que ainda resta para pagar, apresentou F-Fisher 22,03. As demais variáveis apresentaram resultados significativos, porém, de menor importância.

Tabela 3 - Resultado do Teste de Hipóteses F-Fisher para Função Causa Raiz

Variáveis	F-Fisher	Valor-p
Q32 – Motivos que levaram a possuir dívidas em atraso	394,1921	0,0000
Q36 – Ocasões em que o nome foi incluído no sistema de proteção ao crédito	178,9235	0,0000
Q33 – Número de vezes que renegociou dívidas em atraso	64,5169	0,0000
Q5 - Número de Filhos	31,9130	0,0000
Q27 - Saldo a vencer do empréstimo e/ou financiamento	22,0258	0,0000
Q24 – Formas de pagamento do cartão de crédito	18,7081	0,0000
Q35 – De quem o respondente já emprestou o crédito	18,3723	0,0000
Q10 – Número de pessoas dependentes	17,7293	0,0000
Q28 – Número de parcelas que restam para pagar do empréstimo e/ou financiamento	14,6397	0,0000
Q34 – Para quem o respondente já emprestou o crédito	13,5665	0,0000
Q29 – Formas de pagamento quando realiza compras	12,7479	0,0000
Q14 – Tipo de moradia	12,7066	0,0000
Q25 – Valor da parcela do empréstimo e/ou financiamento	12,1378	0,0000
Q2 – Idade	10,3398	0,0000
Q22 – Limite do cartão de crédito	10,2219	0,0000
Q23 – Ocasões em que utiliza o limite da conta corrente	9,4896	0,0000
Q3 - Estado Civil	7,7202	0,0001
Q11 – Percentual da renda gasto com necessidades básicas	6,9317	0,0001
Q38 – Como o respondente cuida da saúde	5,9831	0,0001
Q15 – Tempo que reside no atual endereço	5,6049	0,0000
Q9 – Renda mensal individual	5,5841	0,0002
Q4 – Maior escolaridade completa	5,1131	0,0000
Q20 – Grau de instrução do chefe da família	3,4489	0,0011
Q21 – Tipo de conta bancária	2,3851	0,0267
Q8 - Tempo trabalho no atual emprego	2,0318	0,0713
Q30 – O que influenciou a realizar compras sem necessidade	1,8026	0,1256

Na Figura 1 a seguir, é possível evidenciar os resultados extraídos para o teste da Função Causa Raiz:

Figura 1 - Resultado Visual do Teste da Função Causa Raiz



Dentre as variáveis analisadas foram excluídas as variáveis: Tempo de trabalho; e Compra sem Necessidade, uma vez que elas não influenciam no fato de o respondente ser caracterizado como adimplente ou inadimplente.

5.4 EXCLUSÃO DAS VARIÁVEIS A PARTIR DA ANÁLISE DA MULTICOLINEARIDADE

A multicolinearidade refere-se à forte correlação entre duas ou mais variáveis explicativas incluídas num modelo de regressão. Em razão dessa anomalia, tem-se um aumento da variância da estimativa e, por consequência um aumento do erro padrão de estimativa e uma redução do teste *t-student*, o que acaba induzindo a se aceitar uma hipótese nula quando na verdade ela deveria ser rejeitada.

A Figura 2 apresenta os coeficientes de determinação ou de explicação (R^2) entre as variáveis independentes remanescentes da subseção descrita na análise anterior, visando detectar efeitos de multicolinearidade.

Figura 2 - Detecção da Multicolinearidade Entre as Variáveis Independentes

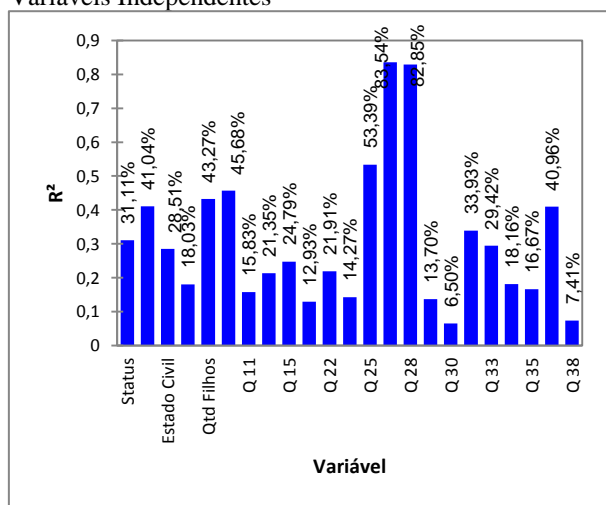
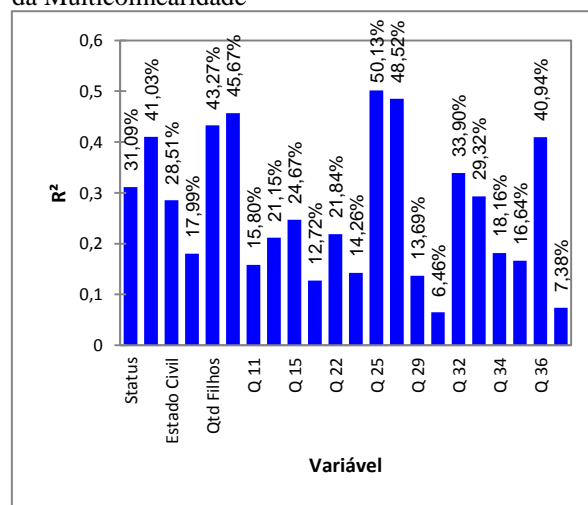


Figura 3 - Variáveis Independentes sem a Presença da Multicolinearidade



Os resultados expressos na Figura 2 mostram que existem multicolinearidade entre as variáveis Q27 (Valor do empréstimo ou financiamento que ainda resta par pagar) e Q28 (Quantidade de parcelas que ainda restam para pagar do empréstimo ou financiamento), com um coeficiente de explicação iguais 83,54% e 82,85% sendo superior a 70% que é o limite definido por Matos (2000), como sendo danoso para a construção do modelo em questão.

Neste sentido, optou-se por excluir a variável Q27 dado que a mesma possui o maior valor de coeficiente de explicação R². A Figura 3, por sua vez, mostra os resultados após a eliminação de Q27. Observe que após a eliminação da variável Q27, não foram vistos os efeitos da multicolinearidade entre as variáveis independentes.

5.5 ESTIMAÇÃO E VALIDAÇÃO DO MODELO DE REGRESSÃO LOGÍSTICA BINÁRIA

Para a construção do modelo de previsão de inadimplência, utilizou-se a técnica de regressão logística binária. Considerando que o percentual de inadimplentes foi igual a 14,14%, fruto da razão existente entre o número de pessoas caracterizadas como inadimplentes (335 indivíduos) e o total de respondentes (2.369 indivíduos) partiu-se dos dados que não foram utilizados na amostra de desenvolvimento e selecionou-se uma amostra aleatória de 200 indivíduos para a amostra de validação, com o objetivo de verificar a aderência do modelo estimado aos dados observados. Esta amostra possibilita comprovar se o modelo está predizendo os dados adequadamente.

Realizou-se a modelagem usando o procedimento de seleção de variáveis *Stepwise (Backward)*, onde se constrói o modelo retirando as variáveis passo a passo. Neste método, o modelo inicial é o mais amplo possível onde são retirados os regressores a cada passo, baseado em quanto eles melhorariam o modelo na medida em que são excluídos. Quando nenhum regressor a mais pode ser excluído, não produzindo uma melhoria substancial, o modelo final é gerado.

Nesse contexto, para o desenvolvimento do modelo estimou-se sucessivamente um total de 17 modelos alternativos, onde foram excluídas as variáveis que não agregassem maior significância estatística aos índices do modelo. Os critérios de parada para a exclusão foram os Critérios de Informação de Akaike e Critério *Bayesiano* de Schwarz. Quanto maiores

forem os seus valores, mais parcimonioso será o modelo estimado. Este procedimento foi realizado no software Xlstat versão demonstração.

A estatística *-2Log Verossimilhança* refere-se a um indicador onde busca aferir a capacidade de o modelo estimar a probabilidade associada à ocorrência do evento, neste caso, a inadimplência. Trata-se de um teste de hipóteses de qui-quadrado, em que indica o potencial dos parâmetros inseridos no modelo visando estimar a probabilidade de ocorrência do evento (CORRAR; PAULO; DIAS FILHO, 2007). Os resultados das estatísticas dos passos da modelagem *logit* binomial, passo-a-passo, encontram-se na Tabela 4, apresentada a seguir:

Tabela 4 -Processo de Escolha do Modelo de Regressão Logística Binária – Método Stepwise (Backward)

V	Variáveis	Variável IN/OUT	Status	-2 Log (Verossimilhança)	Pr > LR	Pr > Score	Pr > Wald	AIC de Akaike	SBC de Schwarz
21	Idade / E.Civil / Escol. / Qtd. Filhos Dep. / Q10 / Q11 / Q14 / Q15 / Q20 / Q22 / Q23 / Q25 / Q28 / Q29 / Q30 / Q32 / Q33 / Q34 / Q35 / Q36 / Q38	-	-	1574,620	0,000	0,000	0,000	1618,620	1743,625
20	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q10 / Q11 / Q14 / Q15 / Q20 / Q22 / Q23 / Q25 / Q28 / Q29 / Q30 / Q32 / Q33 / Q35 / Q36 / Q38	Q34	OUT	1574,622	0,000	0,000	0,000	1618,622	1743,626
19	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q11 / Q14 / Q15 / Q20 / Q22 / Q23 / Q25 / Q28 / Q29 / Q30 / Q32 / Q33 / Q35 / Q36 / Q38	Q10	OUT	1574,628	0,000	0,000	0,000	1618,628	1743,633
18	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q11 / Q14 / Q20 / Q22 / Q23 / Q25 / Q28 / Q29 / Q30 / Q32 / Q33 / Q35 / Q36 / Q38	Q15	OUT	1574,635	0,000	0,000	0,000	1618,635	1743,639
17	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q11 / Q14 / Q20 / Q22 / Q23 / Q25 / Q29 / Q30 / Q32 / Q33 / Q35 / Q36 / Q38	Q28	OUT	1574,656	0,000	0,000	0,000	1618,656	1743,660
16	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q11 / Q14 / Q20 / Q22 / Q23 / Q25 / Q29 / Q30 / Q32 / Q33 / Q36 / Q38	Q35	OUT	1574,682	0,000	0,000	0,000	1618,682	1743,687
15	Idade / E. Civil / Escol. / Qtd. Filhos Dep. / Q14 / Q20 / Q22 / Q23 / Q25 / Q29 / Q30 / Q32 / Q33 / Q36 / Q38	Q11	OUT	1574,705	0,000	0,000	0,000	1618,705	1743,709
14	Idade / E. Civil / Escol. / Q14 / Q20 / Q22 / Q23 / Q25 / Q29 / Q30 / Q32 / Q33 / Q36 / Q38	Qtd. Filhos Dep.	OUT	1574,772	0,000	0,000	0,000	1618,772	1743,776
13	Idade / Escol. / Q14 / Q20 / Q22 / Q23 / Q25 / Q29 / Q30 / Q32 / Q33 / Q36 / Q38	E. Civil	OUT	1574,821	0,000	0,000	0,000	1618,821	1743,825
12	Idade / Escol. / Q14 / Q20 / Q22 / Q23 / Q25 / Q30 / Q32 / Q33 / Q36 / Q38	Q29	OUT	1574,912	0,000	0,000	0,000	1618,912	1743,916
11	Idade / Escol. / Q14 / Q22 / Q23 / Q25 / Q30 / Q32 / Q33 / Q36 / Q38	Q20	OUT	1575,083	0,000	0,000	0,000	1619,083	1744,088
10	Escol. / Q14 / Q22 / Q23 / Q25 / Q30 / Q32 / Q33 / Q36 / Q38	Idade	OUT	1575,312	0,000	0,000	0,000	1619,312	1744,316
9	Escol. / Q14 / Q22 / Q23 / Q25 / Q32 / Q33 / Q36 / Q38	Q30	OUT	1575,526	0,000	0,000	0,000	1619,526	1744,531
8	Escol. / Q14 / Q22 / Q23 / Q32 / Q33 / Q36 / Q38	Q25	OUT	1575,897	0,000	0,000	0,000	1619,897	1744,902
7	Escol. / Q22 / Q23 / Q32 / Q33 / Q36 / Q38	Q14	OUT	1576,313	0,000	0,000	0,000	1620,313	1745,317
6	Q22 / Q23 / Q32 / Q33 / Q36 / Q38	Escol.	OUT	1576,959	0,000	0,000	0,000	1620,959	1745,964
5	Q22 / Q23 / Q32 / Q33 / Q36	Q38	OUT	1578,309	0,000	0,000	0,000	1622,309	1747,313

Observe por meio da Tabela 4 que o melhor modelo a ser estimado possui apenas 5 variáveis explicativas ou independentes. Assim, as variáveis que conseguem estimar o modelo foram retidas pelo modelo de regressão logística, isto é, aquelas que conseguem prever a variável dependente *status* referem-se: Q22: Limite do cartão de crédito; Q23: Ocasões em que utiliza o limite da conta corrente; Q32: Motivos que levaram a possuir dívidas em atraso; Q33: Número de vezes que renegociou dívidas em atraso e Q36: Ocasões em que o nome foi incluído no sistema de proteção ao crédito, além da constante do modelo.

A Tabela 5 evidencia a estimação do modelo de regressão logística binário para as variáveis finais selecionadas pelo método de estimação de *Stepwise (Backward)*.

Tabela 5 - Estimação do Modelo Final de Regressão Logística Binária

Fonte	Parâmetros	Qui-quadrado (Wald)	Pr > Wald	Razão de Odds
Constante	-3,917	52,707	<0,0001	
Q22	-0,068	6,345	0,012	0,925
Q23	0,046	5,349	0,021	1,047
Q32	0,433	70,398	<0,0001	1,541
Q33	0,092	6,529	0,011	1,097
Q36	0,230	30,565	<0,0001	1,258

Observando os resultados dispostos na Tabela 5, pode se evidenciar que existe um relacionamento negativo entre o limite do cartão de crédito e a probabilidade do indivíduo ser caracterizado como inadimplente. Assim, quanto menor é o limite do cartão de crédito maior é a probabilidade dele se tornar inadimplente.

Observou-se ainda que os valores estimados para a estatística do qui-quadrado de verossimilhança (LR) foram todos significantes ao nível de 5%. Todos os valores de probabilidades para os coeficientes estimados situaram-se abaixo de 0,05.

Verifica-se também que existe um relacionamento positivo entre a variável Q23: Ocasões em que utiliza o limite da conta corrente e o fato do indivíduo ser caracterizado como inadimplente. Assim, o resultado denota que o aumento no limite da sua conta corrente reflete-se numa elevação do nível de inadimplência.

Tem-se ainda um relacionamento positivo entre a variável Q32: Motivos que levaram a possuir dívidas em atraso e o fato de o indivíduo ser caracterizado como inadimplente. A partir do valor do coeficiente estimado pode-se afirmar que essa motivação decorre do fato de que o indivíduo possui maior descontrole financeiro ou mesmo falta de planejamento financeiro. Nesse contexto, quanto maior forem tais motivos, maior será também o seu nível de inadimplência.

Observou-se um relacionamento positivo entre a variável Q33: Número de vezes que renegociou dívidas em atraso e o fato do indivíduo ser caracterizado como inadimplente. Neste caso, pode-se deduzir que quanto mais o indivíduo renegocia as suas dívidas, maior é a sua capacidade de compra e, por conseguinte, de incorrer em inadimplência.

Viu-se também um relacionamento positivo entre a variável Q36: Ocasões em que o nome foi incluído no sistema de proteção ao crédito e o fato do indivíduo ser caracterizado como inadimplente. Assim, quanto menor for a quantidade de vezes que os indivíduos têm o nome incluso no SPC menor também será, por consequência, a inadimplência.

A Tabela 6 a seguir, evidencia as estatísticas de bondade de ajustamento do modelo formulado.

Tabela 6 - Estatísticas da Bondade de Ajustamento

Estatística	Independente	Completo
-------------	--------------	----------

Observações	2.169	2.169
Soma dos pesos	2.169	2.169
GL	2.168	2.163
-2 Log (Verossimilhança)	1.793,87	1.574,62
R ² (McFadden)	0,000	0,122
R ² (Cox e Snell)	0,000	0,096
R ² (Nagelkerke)	0,000	0,171
AIC	1.795,87	1.586,62
SBC	1.801,55	1.620,71
Iterações	0	6

Os valores estimados para os pseudos R² da regressão logística foram considerados baixos se comparados com a mesma estatística de uma regressão linear. Apesar de saber que esses valores não podem ser comparados com os resultados obtidos em estudos similares, tendo em vista que essa comparação somente seria possível quando se utilizam a mesma amostra e as mesmas variáveis.

O valor, por exemplo, do coeficiente R² (*Nagelkerke*) para o modelo estimado indica que 17,08% da variância da variável dependente é explicada pelas variáveis independentes do modelo.

O teste de hipóteses de Hosmer e Lemeshow (1989) que avalia se existem diferenças estatisticamente significativas entre as classificações previstas e as observadas, onde divide os casos em classes e compara com as frequências preditas e observadas, em cada classe, por meio da estatística de qui-quadrado, pode ser visto na Tabela 7.

Tabela 7 - Teste de Hipóteses de Hosmer-Lemeshow

Estatística	Qui-quadrado	GL	Pr > Qui ²
Estatística de Hosmer-Lemeshow	110,542	8	< 0,0001

Veja a partir da Tabela 7 que o valor de probabilidade (Pr) para o teste de hipóteses de *Hosmer-Lemeshow* situou-se abaixo do nível de significância estatística adotada (0,05), rejeitando-se com isso, a hipótese nula de que não existem diferenças significativas entre os valores previstos e observados, o que denota que o modelo é capaz de produzir classificações confiáveis.

A Equação estimada para a variável binária *status* pode ser descrita algebricamente da seguinte forma:

$$\hat{Y}(1) = \frac{1}{1 + e^{-Z}}$$

Onde:

$$Z = -3,917 - 0,00783 * Q22 + 0,00462 * Q23 + 0,4326 * Q32 + 0,00923 * Q33 + 0,2298 * Q36$$

Figura 3 – Equação para Variável Binária

A Tabela 8 evidencia a tabela de classificação para a amostra de estimação que examina o percentual de indivíduos classificados correta ou incorretamente.

Tabela 8 - Tabela de Classificação para a Amostra de Estimação

de \ a	0	1	Total	% correto
--------	---	---	-------	-----------

Adimplentes (0)	1.855	0	1.855	100,00%
Inadimplentes (1)	314	0	314	0,00%
Total	2.169	0	2.169	85,52%

Observe por meio da Tabela 8 que o nível de assertividade do modelo estimado foi de 85,52% com 1.855 indivíduos classificados corretamente como adimplentes enquanto que 314 foram classificados como inadimplentes. Esse índice de acerto da amostra de estimação do modelo logístico apresentou resultados considerados como bons em termos de classificação, já que nas visões de Pcinini, Oliveira e Monteiro (2003), bem como Selau e Ribeiro (2009), os modelos de *credit scoring* com taxas de acertos acima de 65% são considerados como bons modelos aos olhos dos especialistas em modelagem desta natureza.

A Tabela 9 mostra os resultados das projeções para a amostra de validação do modelo de regressão logística estimado, considerando a separação de 200 casos para se avaliar essa *performance*.

Tabela 9 - Tabela de Classificação para a Amostra de Validação

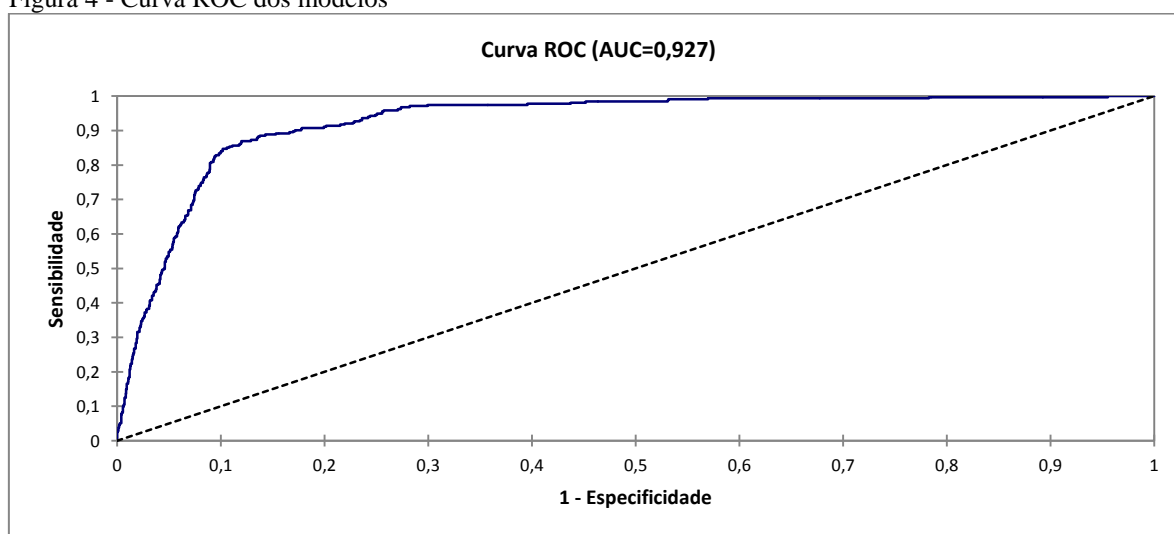
de \ a	0	1	Total	% correto
Adimplentes (0)	174	0	174	100,00%
Inadimplentes (1)	26	0	26	0,00%
Total	200	0	200	87,00%

Observe que para a amostra de validação, o modelo estimado consegue classificar corretamente 87% dos casos. Veja que 174 indivíduos foram classificados corretamente como adimplentes enquanto que 26 indivíduos foram classificados como inadimplentes.

Notou-se ainda que a assertividade dos indicadores avaliados anteriormente está em conformidade com os estudos de Casa Nova (2013), Camargos, Araújo e Camargos (2012), Selau e Ribeiro (2009), Palmuti e Picchiali (2012) e Ferreira, Celso e Barbosa Neto (2012) onde os estudos realizados se assemelham a este trabalho.

A Figura 4 mostra o gráfico da curva ROC que, permite aferir a acurácia dos modelos utilizados na classificação do risco de inadimplência.

Figura 4 - Curva ROC dos modelos



Verifica-se com base na Figura 4 que a área abaixo da curva ROC é igual a 0,927 e que, ao nível de significância estatística de 5%, existem evidências que levam a rejeição da hipótese nula de seja igual a 0,5 (área abaixo da reta que retrata o poder de discriminação nulo), ou seja, o modelo estimado apresenta um excelente poder discriminatório.

6 CONSIDERAÇÕES FINAIS

A utilização dos modelos de *Credit Scoring* é importante para as organizações, pois possibilita tomar decisões de forma mais rápida e segura, possibilitando reduzir os riscos e o nível de inadimplência, além de agregar uma gama de informações relativas à forma como as características do solicitante do crédito e do seu negócio afetam na inadimplência.

A presente pesquisa analisou o nível de inadimplência das pessoas físicas de baixa renda da cidade de Curitiba e região metropolitana e a partir da resolução 2.682/99 do Banco Central do Brasil, estimou um modelo de *Credit Score* que melhor segmenta os adimplentes e inadimplentes, pessoas físicas de baixa renda da cidade de Curitiba e região metropolitana.

Para classificar os clientes em bom ou mau pagadores foi adotada como base a Resolução 2682/99 do Banco Central e ficou convencionado que bom cliente seria aquele com atraso até 90 dias e mau pagador ou inadimplentes os respondentes que declararam possuir dívidas com atraso superior a 90 dias. Mediante os resultados obtidos após a classificação dos clientes em bom ou mau pagador observou-se que apenas 335 são considerados maus pagadores, por possuírem três parcelas ou mais em atraso o que corresponde ao prazo acima de 90 dias.

Quando realizada a análise exploratória de cada variável versus o status de bom ou mau pagador pode-se observar a partir do qui quadrado que das 37 variáveis independentes 9 não apresentaram resultados significativos, 4 por possuírem múltiplas respostas não foi possível analisar o teste qui quadrado e 24 variáveis apresentaram resultados significantes. Em geral os resultados encontrados foram satisfatórios, o que possibilita observar por meio da Curva ROC a qual mostra que o modelo estimado apresenta um excelente poder discriminatório.

Quando realizada a estimação e validação do modelo de regressão logística binária para o desenvolvimento do modelo, estimou-se sucessivamente um total de 17 modelos alternativos, onde foram excluídas as variáveis que não agregassem maior significância estatística aos índices do modelo e o melhor modelo a ser estimado possui apenas 5 variáveis explicativas ou independentes.

Por fim, após qualificada análise dos dados, se pode afirmar que as variáveis passíveis de estimar o modelo foram retidas pelo modelo de regressão logística, isto é, aquelas que conseguem predizer a variável dependente *status* são: Q22: Limite do cartão de crédito; Q23: Ocasões em que utiliza o limite da conta corrente; Q32: Motivos que levaram a possuir dívidas em atraso; Q33: Número de vezes que renegociou dívidas em atraso e Q36: Ocasões em que o nome foi incluído no sistema de proteção ao crédito, além da constante do modelo.

REFERÊNCIAS

- ABEP – Associação Brasileira de Empresas de Pesquisa. Critério de Classificação Econômica Brasil. Disponível em: <<http://www.abep.org/novo/FileGenerate.ashx?id=297>>. Acesso em: 11 fev. 2013.
- ANDRICH, E. G.; CRUZ, J. A. W.; ANDRICH, R. G.; GUINDANI, R. A. Finanças corporativas: análise de demonstrativos contábeis e de investimentos. Editora Intersaberes, 2014.

- EVERY, R. B.; BREVOORT, K. P.; CANNER, G. B. Credit scoring and its effects on the availability and affordability of credit. *The Journal of Consumer Affairs*, v. 43, n. 3, p. 380-542, Fall 2009.
- BANCO CENTRAL DO BRASIL. Resolução nº 2682/99. Disponível em: http://www.bcb.gov.br/pre/normativos/res/1999/pdf/res_2668_O.pdf. Acesso em: 01 abr. 2013.
- BRUSKY, B.; FORTUNA, J. P. Entendendo a demanda para as microfinanças no Brasil: um estudo de caso qualitativo de duas cidades. Programa de Desenvolvimento Institucional do Banco Nacional de Desenvolvimento (PDI/BNDES), (Relatório de Pesquisa), 2002.
- BUENO, V. F. F. Avaliação de risco na concessão de crédito bancário para micros e pequenas empresas. Florianópolis: UFSC, 2003. Dissertação (Mestrado em Engenharia de Produção), Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, 2003.
- CAMARGOS, M. A.; ARAÚJO, E. A. T.; CAMARGOS, M. C. S. A inadimplência em um programa de crédito de uma instituição financeira pública de Minas Gerais: uma análise utilizando regressão logística. *REGE*, v. 19, n. 3, p. 473-492, jul./set. 2012 - São Paulo – SP, Brasil.
- CAOQUETTE, J. B.; ALTMAN, E. I.; NARAYANAN, P. Gestão do Risco de Crédito: o grande desafio dos mercados financeiros globais. 2. Ed. Rio de Janeiro: Qualitymark, SERASA, 2009.
- CASA NOVA, S. P. C. Quanto pior, melhor : Estudo da utilização da análise por envoltória de dados em modelos de análise de inadimplência/insolvência de empresas. *Revista Contemporânea de Contabilidade*, v. 10, n. 19, p. 71–96. 2013.
- CASADO, M. M. Os princípios fundamentais como ponto de partida para uma primeira análise do sobre endividamento no Brasil. *Revista de Direito do Consumidor*. São Paulo: RT, n.33, 2000.
- CORRAR, L.J.; PAULO, E.; DIAS FILHO, J.M. *Análise Multivariada*. 1 ed. São Paulo: Atlas, 2007.
- FERREIRA, M. A. M.; CELSO, A. S. S.; BARBOSA NETO, J. E. Aplicação do modelo logit binomial na análise do risco de crédito em uma instituição bancária. *Revista de Negócios*, v.17, n. 1, p. 41–59. 2012.
- GHODSELAHI, A. A hybrid support vector machine ensemble model for credit scoring. *International Journal of Computer Applications*, v. 17, n. 5, p. 975-8887, 2011.
- LOUZADA, F. et al. Statistical Measured of the Predictive Capacity of Classification models used in Credit Scoring. *Technology Credit - Serasa Experian*, v. 68. São Paulo, 2009.
- MATTOZO, C. Q. M. Identidade, inserção social e acesso a serviços financeiros: um estudo na favela da Rocinha. Rio de Janeiro: CPPEAD/UFRR, apresentado com Tese de Doutorado, UFRJ, 2005.
- MOREIRA, F. A. A. Demanda e oferta de entretenimento: um estudo do segmento de baixa renda do Distrito de Itaquera na cidade de São Paulo. 2006. 162 p. Dissertação (Mestrado em Administração) – Escola de Administração de Empresas de São Paulo, São Paulo, 2006.
- NOBRE, L. H.; MACEDO, A. F. P.; NOBRE, F. C.; SILVA, W. V. Análise da Relação entre Variáveis Demográficas e Escores de Tolerância ao Risco. In: *Revista de Administração da UFSM*, v. 10, n. 1, 2017.
- PALMUTI, C. S.; PICCHIAI, D. Mensuração do risco de crédito por meio de análise estatística multivariada. *Revista Economia Ensaios*, v. 26, n. 2, p. 7–22. 2012.
- RIBEIRO, C. F.; ZANI, J.; ZANINI, F. A. M. Estimativa da verificação da probabilidade de inadimplência: uma verificação empírica na Universidade Católica de Pelotas. In: *ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS GRADUAÇÃO E PESQUISA EM*

- ADMINISTRAÇÃO, 33, 2009, São Paulo, Anais. Rio de Janeiro: ANPAD, 2009.
- ROCHA, D. T.; MARTINS, T. S.; SILVA, W. V.; CRUZ, J. A. W. C. Análise de risco: um estudo bibliométrico e sociométrico da produção científica da área de finanças do EnANPAD 1997-2008. In: Revista Pensar Contábil, Vol. 12 (47), 2010.
- SANTOS, J. O. Análise de Crédito: empresas e pessoas físicas. São Paulo: Atlas, 2009.
- SECURATO, J. R.; PEROBELLI, F. F. C. Comparação entre métodos para determinação do valor presente de uma carteira de crédito e de seu risco. Cadernos de Pesquisa em Administração. v. 7, n. 4, out./ dez.2000.
- SELAU, L. P. R.; RIBEIRO, J. L. D. Uma sistemática para construção e escolha de modelos de previsão de risco de crédito. Revista Gestão e Produção, v. 16, n. 3, p. 398–413. 2009.
- SEMEDO, DANILSON, P.V. Credit Scoring: Aplicação da regressão logística vs redes neurais artificiais na avaliação do risco de crédito no mercado Cabo-Verdiano. 2009. 126 p. Dissertação (Mestrado em Estatística e Gestão da Informação) – Instituto Superior de Estatística e Gestão da Informação da Universidade Nova de Lisboa.
- SICSÚ, A. L. Credit Scoring: Desenvolvimento-Implantação-Acompanhamento. 1.Ed. São Paulo: Blucher, 2010.
- SILVA, J. S. S. Aspectos do superendividamento do consumidor de baixa renda. 20 p. Escola de Magistratura do Estado do Rio de Janeiro, Rio de Janeiro, 2012.
- SILVA, J. P. Gestão e Análise de risco de Crédito. 5. Ed. São Paulo: Atlas, 2006.
- THOMAS, L. A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. International Journal of Forecasting, Amesterdam, v.16, n.2, Apr/Jun., 2000.