

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO DE COOPERATIVAS  
MESTRADO EM GESTÃO DE COOPERATIVAS  
ESCOLA DE NEGÓCIOS**

**DANILO IANCKIEVICZ VASCONCELLOS**

**MODELAGEM DE RISCO DE CRÉDITO: UM ESTUDO DO SEGMENTO DE  
PESSOAS FÍSICAS EM UMA COOPERATIVA DE CRÉDITO DO PARANÁ**

**CURITIBA**

**2018**

**DANILO IANCKIEVICZ VASCONCELLOS**

**MODELAGEM DE RISCO DE CRÉDITO: UM ESTUDO DO SEGMENTO DE  
PESSOAS FÍSICAS EM UMA COOPERATIVA DE CRÉDITO DO PARANÁ**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão de Cooperativas, da Escola de Negócios, da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de Mestre em Gestão de Cooperativas.

Orientador: Prof. Dr. Pedro Guilherme Ribeiro Piccoli

**CURITIBA**

**2018**

*Dedico este trabalho à minha mãe Ana Cristina, ao meu pai David, à minha irmã Júlia, à minha madrinha Maria, às minhas avós Tereza e Sebastiana e a todos que estiverem ao meu lado nesse importante período de minha vida.*

## AGRADECIMENTOS

Aos meus pais, pela estrutura, esforço e ensinamentos que sempre dedicaram a mim.

À minha madrinha e minhas avós, pelas palavras de incentivo, cobrança e perseverança.

A cooperativa X e aos meus gestores, pela oportunidade de realizar o mestrado em conjunto do serviço, no período em que lá estive. Além de colegas de serviço, pelas conversas, *feedbacks* e discussões - “no bom sentido” – que tivemos nesse período juntos.

Aos professores do Mestrado em Gestão de Cooperativas da PUCPR pela aprendizagem, pelo convívio e pelos incentivos.

Ao Prof. Dr. Pedro Guilherme Ribeiro Piccoli, meu orientador, pelas orientações, conversas e experiências passadas. E também, pela paciência, empatia e camaradagem durante todo esse ciclo.

Aos colegas de curso, e agora amigos que conquistei, pelos mais variados momentos, dos mais prazerosos aos de maiores dificuldades, e pelas mais diversas experiências, das aulas e discussões até a nossa viagem à Alemanha.

Ao Sistema Ocepar, em especial à equipe do SESCOOP, pelo apoio financeiro.

À Pontifícia Universidade Católica do Paraná e seus funcionários, por toda a estrutura e auxílio prestado.

E por fim, mas não menos importante que todos os demais, à Deus por ter me dado todas as condições necessárias.

## RESUMO

Diante do crescimento do mercado de crédito no Brasil, da expansão do cooperativismo nesse mercado e da cooperativa X estar convergindo nesta mesma direção, o presente estudo objetivou estimar e apresentar duas opções de modelos de *credit score* que contribua para o aumento de automação no processo de análise e concessão de crédito, aos cooperados pessoas físicas, na cooperativa X. Para tal, foi disponibilizado uma base de dados, com o total de 6130 solicitações de crédito do período de 2013 até 2015, com informações cadastrais, comportamento e restrições de mercado. Tais informações foram empregadas na regressão logística binária, afim de gerar o modelo estatístico que auxilie o processo de automação na análise e concessão de crédito dos solicitantes futuros. Foi-se utilizado os *softwares* MS Excel 2016 e *Data Analysis and Stastiscal Software* (STATA) para sumarização e aplicação dos métodos estatísticos. Metodologicamente, este estudo foi definido como quantitativo em relação a sua natureza e explicativo quanto ao tipo. Dentre as informações disponibilizadas, foram selecionadas 20 variáveis que traziam com fidedignidade as informações. As variáveis que foram selecionadas na regressão logística binária para pelo menos um dos modelos foram: Categoria da linha de crédito 1 (Crédito de antecipação de garantias futuras); Cheques devolvidos por insuficiência de saldo no últimos 12 meses; Idade; Média de saldo em conta corrente dos últimos 6 meses; Quantidade de parcelas mensais; Saldo devedor de empréstimos na cooperativa; Consulta Serasa; Tempo de relacionamento; Tipo de conta; Valor de limite de cheque especial; Valor em aplicações financeiras; e valor solicitado. Os modelos apresentaram curva ROC de 0,829 e 0,812, um excelente valor para acurácia dos mesmos. O trabalho decidiu por não apresentar um único percentual de precisão preditiva, tendo em vista que isso variaria de acordo com a definição do *cutoff*, este que deve ser definido de acordo as políticas da cooperativa e seu staff. Entretanto, foi apresentado três simulações em cada modelo, dentre as quais é possível verificar que todas apresentariam um crescimento considerável de automatização da análise e concessão de crédito, se implementado o modelo.

**Palavras-chave:** Cooperativa de crédito. *Credit score*. *Credit score* em cooperativa de crédito. Risco de crédito.

## ABSTRACT

Given the growth of the credit market in Brazil, the expansion of cooperativism in this market and the cooperative X are converging in the same direction. The present study aimed to estimate and present two options of credit score models that contribute to the increase of automation in the process of analysis and granting of credit to individual cooperatives in the cooperative X. For this purpose, a database was made available, with a total of 6130 credit applications from the period 2013 to 2015, with registration information, behavior and market restrictions. Such information was used in the binary logistic regression, in order to generate the statistical model that helps the process of automation in the analysis and credit granting of the future applicants. We used the MS Excel and Data Analysis and Statistical Software (STATA) software to summarize and apply statistical methods. Methodologically, this study was defined as quantitative in nature and explanatory of type. Among the information available, we selected 20 variables that reliably brought the information. The variables that were selected in the binary logistic regression for at least one of the models were: Category of the line of credit 1 (Credit of anticipation of future guarantees); Checks returned for insufficient balance in the last 12 months; Age; Average balance in current account for the last 6 months; Quantity of monthly installments; Balance borrowed from the cooperative; Consultation Serasa; Relationship time; Account Type; Overdraft limit value; Amount in financial investments; and requested amount. The models presented a ROC curve of 0.829 and 0.812, an excellent value for their accuracy. The work decided not to present a single percentage of predictive accuracy, considering that this would vary according to the cutoff definition, which should be defined according to the policies of the cooperative and its staff. However, three simulations were presented in each model, among which it is possible to verify that all would present a considerable growth of automation of the analysis and credit concession if the model was implemented.

**Key-words:** Credit union. Credit score. Credit score in credit union. Credit risk.

## LISTA DE FIGURAS

Figura 1 - Crédito\PIB segundo origem de capital – Brasil (1995-2011) .....	12
Figura 2 - Número cooperativas de crédito .....	18
Figura 3 - Etapas do desenvolvimento do <i>credit score</i> .....	29
Figura 4 - Histograma da variável valor renda.....	43
Figura 5 - Histograma da variável LN valor renda.....	44
Figura 6 - ROC do modelo 1 .....	50
Figura 7 - ROC do modelo 2 .....	57

## LISTA DE QUADROS

Quadro 1 - <i>Rating</i> \percentual de provisionamento .....	28
Quadro 2 - Variáveis e frequências .....	31
Quadro 3 - Linhas de crédito por categorias .....	35
Quadro 4 - Variáveis .....	38



## LISTA DE TABELAS

Tabela 1 - Participação do cooperativismo de crédito no sistema financeiro nacional .....	19
Tabela 2 - Propostas, categorias e inadimplência 1 .....	34
Tabela 3 - Propostas, categorias e inadimplência 2 .....	34
Tabela 4 - Variáveis e coeficientes do modelo final 1 .....	45
Tabela 5 - Matriz correlação do modelo 1 .....	48
Tabela 6 - VIF do modelo 1 .....	48
Tabela 7 - Teste <i>Hosmer-Lemeshow</i> do modelo 1 .....	49
Tabela 8 - Matriz de classificação do modelo 1 .....	51
Tabela 9 - Percentual de aprovações automáticas .....	52
Tabela 10 - Variáveis e coeficientes do modelo final 2 .....	53
Tabela 11 - Matriz correlação do modelo 2 .....	55
Tabela 12 - VIF modelo 2 .....	55
Tabela 13 - Teste Hosmer-Lemeshow do modelo 2 .....	56
Tabela 14 - Matriz de classificação do modelo 2 .....	58
Tabela 15 - Percentual de Aprovações Automáticas 2 .....	58

## LISTA DE ABREVIATURAS E SIGLAS

BACEN	- Banco Central do Brasil
BCB	- Banco Central do Brasil
BCBC	- <i>Basel Committee on Banking Supervision</i>
CPF	- Cadastro Pessoa Física
IF	- Instituição Financeira
OCB	- Organização das Cooperativas Brasileiras
PIB	- Produto interno bruto
STATA	- <i>Data Analysis and Stastiscal Software</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>11</b>
1.1 APRESENTAÇÃO DO TEMA.....	11
1.2 PROBLEMA DA PESQUISA .....	13
1.3 DEFINIÇÃO DO OBJETIVO GERAL.....	14
1.4 DEFINIÇÃO DOS OBJETIVOS ESPECÍFICOS .....	15
1.5 JUSTIFICATIVAS DO ESTUDO.....	15
<b>2 REFERENCIAL TEÓRICO</b> .....	<b>16</b>
2.1 COOPERATIVISMO DE CRÉDITO NO BRASIL.....	16
2.2 CRÉDITO .....	20
2.3 RISCO DE CRÉDITO.....	22
2.4 <i>CREDIT SCORING</i> .....	27
<b>3 PROCEDIMENTOS METODOLÓGICOS</b> .....	<b>32</b>
3.1 ESPECIFICAÇÃO DO PROBLEMA .....	32
3.2 CARACTERIZAÇÃO DA PESQUISA .....	32
3.3 COLETA E TRATAMENTO DE DADOS .....	33
<b>3.3.1 Dados e produtos de crédito</b> .....	<b>33</b>
<b>3.3.2 Variável resposta</b> .....	<b>36</b>
<b>3.3.3 Variáveis explicativas</b> .....	<b>37</b>
3.4 <i>SOFTWARES</i> UTILIZADOS.....	39
3.5 MÉTODO DE ANÁLISE.....	39
<b>3.5.1 Regressão logística</b> .....	<b>39</b>
<b>3.5.2 Avaliações dos modelos</b> .....	<b>40</b>
<b>4 APRESENTAÇÃO E ANÁLISE DE RESULTADOS</b> .....	<b>43</b>
4.1 MODELOS FINAIS.....	43
<b>4.1.1 Modelo com definição de inadimplência com atraso maior que 30 dias</b> .....	<b>44</b>
<b>4.1.2 Medidas de desempenho do modelo um</b> .....	<b>49</b>
<b>4.1.3 Modelo com definição de inadimplência com atraso maior que 90 dias</b> .....	<b>52</b>
<b>4.1.4 Medidas de desempenho do modelo dois</b> .....	<b>55</b>
<b>5 CONCLUSÃO</b> .....	<b>60</b>
<b>REFERÊNCIAS</b> .....	<b>63</b>

## 1 INTRODUÇÃO

O presente capítulo tem por finalidade apresentar o tema do trabalho em seu contexto geral, o problema de pesquisa, o objetivo geral e os específicos e a justificativa do estudo.

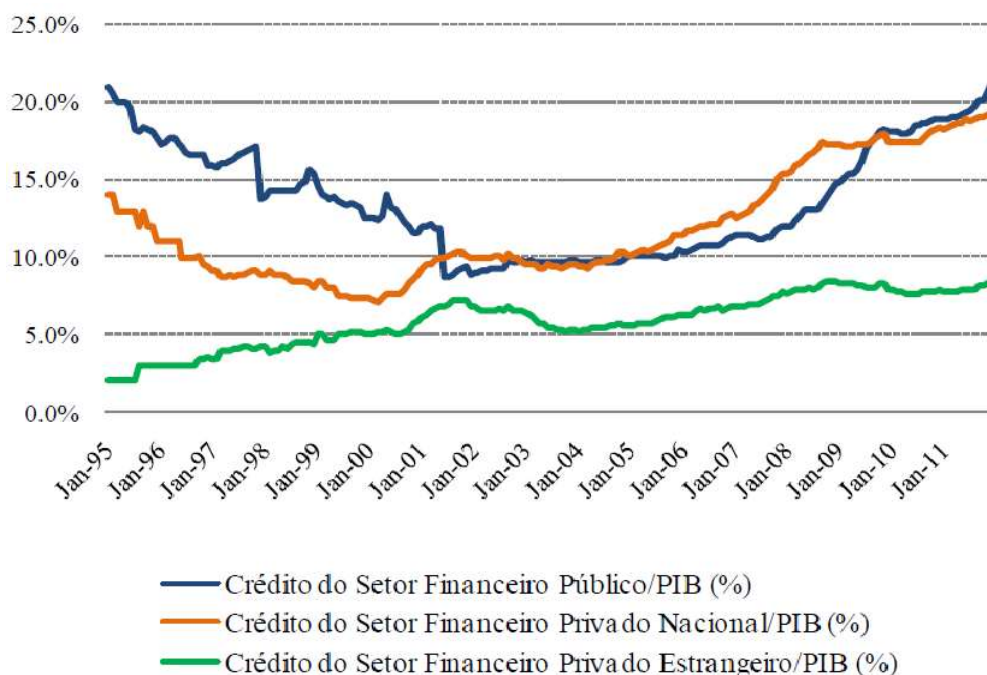
### 1.1 APRESENTAÇÃO DO TEMA

Em diversas literaturas nacionais foi utilizada a premissa de que o mercado de crédito no Brasil teve seu crescimento com a estabilização da moeda. Após o período inflacionário os bancos não obtinham rentabilidade tão facilmente com o *overnight*, e na busca de manter seus ganhos e rentabilidade, viu na intermediação financeira a solução. Segundo Rosa (2000, p. 1), “os empréstimos financeiros passaram a ser um bom negócio para as instituições financeiras pois os mesmos não obtinham mais os grandes lucros vindos da desvalorização da moeda nacional”.

A partir desse contexto, as instituições financeiras almejavam o aumento da carteira de crédito. Contudo, em um primeiro momento, houve uma política monetária restritiva por parte do governo, justamente para reter esses agentes na expansão do crédito, evitando assim a desvalorização da moeda motivada pela inflação de demanda (ANDRADE, 2012). Com a ação restritiva, o crédito teve uma redução de volume, em comparação ao PIB do país, nesses anos subsequentes a estabilização da moeda, conforme a Figura 1.

É possível ainda constatar que foi a partir de 2003 que começou, de forma expressiva, a expansão do crédito por meio das instituições financeiras. Todas essas instituições (públicas, privadas e estrangeiras) somadas, os créditos em relação ao PIB, saltaram de 24% para 47% no período de 2003 até 2011. Costa (2009) afirma que o período em questão foi marcado pelas mudanças institucionais que facilitaram o acesso popular a bancos (através dos correspondentes não-bancários e de programa de conta simplificada) e ao crédito (em consignação, aos consumidores e microcrédito).

Figura 1 - Crédito\PIB segundo origem de capital – Brasil (1995-2011)



Fonte: ANDRADE, 2012, p. 25.

Além disso, houve uma solidificação da macroeconomia no país, acarretando na derrubada da taxa Selic (26,5% a.a. em abril de 2003 para 7,25% a.a. em outubro de 2012). Novamente, devido a essa situação, as instituições financeiras voltaram ainda mais seus olhos para a concessão de crédito, pensando em sua ampliação de rentabilidade.

Com esse amadurecimento da economia brasileira, dos fatores apresentados, o crédito ao consumidor foi de mais fácil acesso e de expansão constante. Paralelamente a estes aspectos, o setor cooperativista de crédito foi conquistando certa notoriedade no cenário financeiro nacional ao longo dos anos. Sua natureza é de certa forma diferente de outras instituições financeiras, tendo em vista que o usuário é também o proprietário dessas instituições.

Essa expansão do crédito na economia do país vem acompanhada de uma maior exposição aos riscos para o sistema financeiro nacional e suas instituições. Tendo em vista, os riscos englobados no funcionamento das instituições financeiras, o risco de crédito é o que está acompanhando em maior grau essa expansão. Ao conceder crédito no mercado, as instituições financeiras estão expostas a não receber futuramente o valor acertado com o agente captador, e quando isso ocorre em escala significativa acaba por colocar não só a instituição em prejuízo, como também, seus demais clientes, acionistas, governo e todo o sistema financeiro do país.

Há regulamentações para as instituições financeiras não se exporem ao risco de crédito de forma disseminada. Porém, as próprias instituições financeiras tendem a aprimorar cada vez mais a mitigação ao risco de crédito, até pela sua própria continuidade no sistema. Um processo utilizado para a mitigação desse risco nas concessões de crédito é o *credit score*, que teve seus primeiros modelos por volta da década de 50. Teve como base os estudos de Fisher (1936) sobre função discriminante e David Durant (1941) que foi precursor em utilizar a técnica para a classificação de clientes em maus ou bons pagadores, motivado também pelo grande volume de analistas de crédito, que começaram a ser convocados para participar da segunda guerra mundial, em conjunto com a criação e expansão do uso do cartão de crédito. Nesse contexto surgiu o modelo de *credit score*, utilizado até nos dias de hoje com evoluções ao longo do tempo.

## 1.2 PROBLEMA DA PESQUISA

O setor do cooperativismo de crédito vem conquistando espaço no mercado financeiro nacional ano a ano (BACEN, 2015) e, em conjunto com a expansão do crédito das últimas décadas, acarretou em um aumento substancial no número de solicitações de crédito por parte dos cooperados na cooperativa de crédito estudada. Para que o crédito não seja concedido de forma disseminada, são realizadas análises econômicas-financeiras sob o cooperado solicitante. Essas análises são de responsabilidade do departamento de crédito, onde se tem a incumbência de aprovar ou negar essas solicitações, utilizando as análises pessoais, por parte dos funcionários do departamento, como forma de discriminação. Esse procedimento ocorre unitariamente sob grande parcela das solicitações, por volta de 80% das solicitações, assim sendo, a baixa automação deste processo acarreta em fatores desfavoráveis como um todo para a cooperativa. Por exemplo, maior custo operacional, tendo em vista o alto custo de mão de obra no país, assim como, a dificuldade de encontrar mão de obra qualificada, ponto este, atentado por Altman, Caoquette e Narayanan (1998, p. 26):

[...] a análise clássica do crédito apresenta diversos problemas, sendo o principal o grande custo de manutenção decorrente da necessidade constante de ter profissionais em treinamento para que estes possam se tornar especialistas”.

Outro ponto negativo, no baixo volume de automação, é a menor agilidade na análise e liberação de crédito, deixando de ofertar uma maior conveniência para o solicitante do crédito, assim como, causando prejuízo para com a diferenciação de prazos, na liberação de crédito, com as demais instituições financeiras concorrentes no mercado. E também, podendo afluir em alta subjetividade por parte dos analistas responsáveis pela análise das solicitações, assim como, diferenciações críticas dos critérios de julgamento sobre os solicitantes, podendo ocasionar controvérsias no senso de justiça, por exemplo, de dois solicitantes de crédito com risco semelhantes quantitativamente, e ocorrendo em diferenciação na aprovação do crédito para estes.

Com este contexto apresentado, a pesquisa tem como escopo buscar o modelo de *credit score* que melhor suplemente e aperfeiçoe as necessidades da cooperativa para a automação na análise e a liberação de crédito junto a pessoas físicas. Considerando isto, busca-se responder ao problema de pesquisa:

**Como estimar um modelo de *credit score*, para solicitantes de crédito pessoas físicas, na cooperativa de crédito estudada, que acarrete em um aumento de automação na análise e concessão de crédito?**

A partir da resposta dessa pergunta de pesquisa, espera-se que o resultado traga ganhos monetários reais para a cooperativa, igualmente como um princípio de paridade futura no modelo e eficácia nas análises e concessões de crédito, em comparação as demais instituições financeiras nacionais, suas concorrentes. Espera-se, também, uma contribuição para a literatura nacional sobre o tema, em ênfase especial, no meio cooperativista de crédito.

### 1.3 DEFINIÇÃO DO OBJETIVO GERAL

Estimar um modelo de *credit score*, para a análise e concessão de crédito automática às pessoas físicas, cooperados em uma cooperativa de crédito do Paraná, que acarrete em um aumento de automação desse processo.

#### 1.4 DEFINIÇÃO DOS OBJETIVOS ESPECÍFICOS

Os objetivos específicos apresentam-se como o desdobramento operacional do objetivo geral do estudo, conforme descritos a seguir:

- a) selecionar dentre todas as informações disponíveis, as variáveis a serem empregadas na modelagem estatística de *credit score*;
- b) constatar as variáveis de maior relevância para a classificação de futuros solicitantes de crédito, em adimplentes ou inadimplentes; e
- c) avaliar a eficiência do modelo desenvolvido, identificando se a aplicação do modelo suplementaria as necessidades para automação na análise e liberação de crédito na cooperativa estudada.

#### 1.5 JUSTIFICATIVAS DO ESTUDO

A pesquisa é justificada em duas perspectivas, a acadêmica e a empírica. Justifica-se academicamente sua relevância, devido, até o presente momento, haver um certo óbice de pesquisa na literatura nacional sobre o tema no meio cooperativista. Possui-se literatura extensa no tema, porém abordada nas demais instituições financeiras, e não se considerando o setor cooperativista do sistema financeiro nacional, o qual pode apresentar peculiaridades ou até mesmo significativas diferenciações de resultados. Empiricamente justifica-se a pesquisa pela ausência de aperfeiçoamento na análise e concessão de crédito automática na cooperativa estudada, e a falta de paridade com seus concorrentes de mercado, ao não se utilizar uma ferramenta estatística, como o *credit score*, que pode estar ocasionando fatores desfavoráveis para concorrência no negócio que a engloba.



## 2 REFERENCIAL TEÓRICO

Busca-se apresentar neste capítulo os principais elementos teóricos para a compreensão do problema empírico estudado, bem como para o alcance dos objetivos propostos. Inicialmente, é realizada uma abordagem introdutória ao cooperativismo de crédito, sua origem e seu contexto no cenário nacional. Posteriormente é descrito fatos sobre o crédito, seus conceitos, história e relevância, indo além, apresenta-se o risco de crédito, seu conceito, sua importância e sua regulamentação. Por fim, é evidenciado os aspectos do *credit score*.

### 2.1 COOPERATIVISMO DE CRÉDITO NO BRASIL

No Brasil, a cultura da cooperação é observada desde a época da colonização portuguesa, esse processo emergiu no Movimento Cooperativista Brasileiro surgido no final do século XIX, estimulado por funcionários públicos, militares, profissionais liberais e operários, para atender às suas necessidades. Em 1902, surgiram as cooperativas de crédito no Rio Grande do Sul, por iniciativa do padre suíço Theodor Amstadt. A partir de 1906, nasceram e se desenvolveram as cooperativas no meio rural, idealizadas por produtores agropecuários muitos deles de origem alemã e italiana. Os imigrantes trouxeram de seus países de origem a bagagem cultural, o trabalho associativo e a experiência de atividades familiares comunitárias, que os motivaram a organizar-se em cooperativas. Com a propagação da doutrina cooperativista, as cooperativas tiveram sua expansão num modelo autônomo, voltado para suprir as necessidades dos próprios membros e assim se livrarem da dependência dos especuladores (OCB-AL, s.d.).

O regime jurídico das sociedades cooperativas no Brasil e a política nacional do cooperativismo foram regulamentados a partir da Lei nº 5.764, de 16 de dezembro de 1971, desde a regulamentação é possível diferenciar, juridicamente, as sociedades cooperativas das sociedades de capital, segundo Bialoskorski Neto (2014):

As decisões de propriedade, como o voto, quórum de assembleia geral e distribuição dos resultados chamados de “sobras” na cooperativa, são baseadas na pessoa cooperada, diferentemente das sociedades de capital, em que essas ações são proporcionais ao capital que cada investidor tem aplicado na organização.

No Brasil, o setor de crédito sempre foi um mercado onde predominaram-se os bancos, estes estão presentes, principalmente, em regiões com maior concentração de pessoas e volume de negócios. Regiões menos populosas e com economias mais rudimentares ficam desassistidas por parte desses agentes, os bancos, isso gera um obstáculo a ser superado para o desenvolvimento econômico e financeiro dessas regiões. Segundo Relatório de Inclusão Financeira (BACEN, 2015), em 2015, 35% dos municípios brasileiros não possuíam postos de atendimento com agências.

Essa deficiência vem sendo suprida com o cooperativismo de crédito, que vem se instalando nessas regiões. Por vezes, passam a ser a primeira e a única instituição financeira a possuir uma agência no local, como também oferecem os mais variados produtos e serviços para os associados. Segundo, novamente, o Relatório de Inclusão Financeira (BACEN, 2015), de 2014 para 2015, houve um aumento de 3% na presença das cooperativas no total de cidades brasileiras.

Quanto à participação no volume de crédito, as demais instituições financeiras, exceto cooperativas de crédito, cresceram, de 2010-2014, 82% no volume, enquanto que as cooperativas de crédito alcançaram um crescimento de 105%, superando em 23% o aumento desse volume no mercado em comparação com as demais instituições (BACEN, 2015).

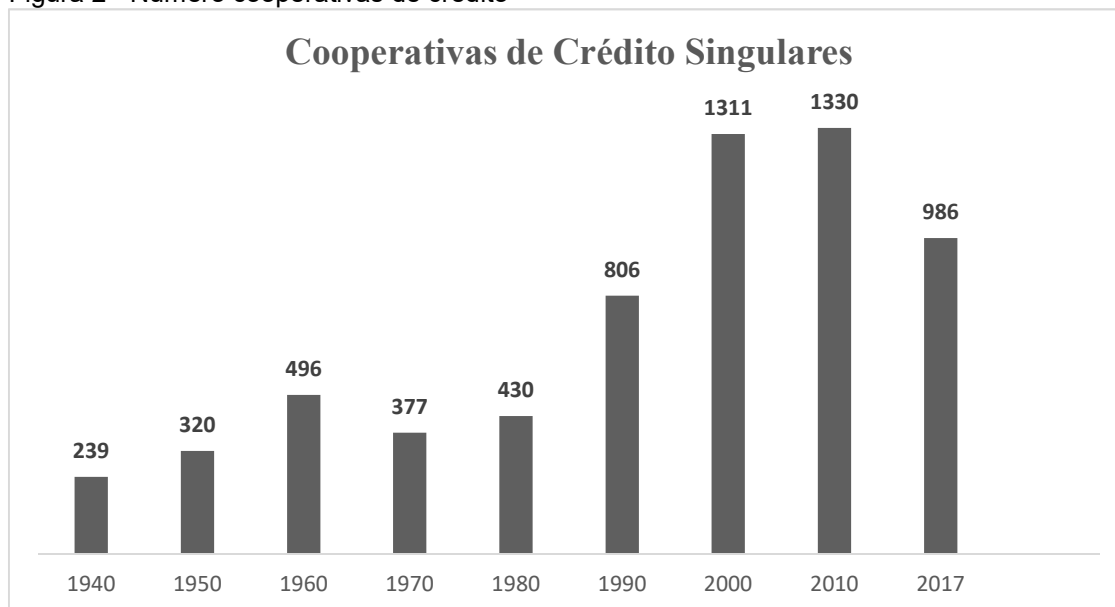
Em números de cooperados, houve um aumento de 27%, de 2012 até 2015, passando de 6 milhões de cooperados em 2012, para mais de 8,3 milhões de cooperados ao fim de 2015. Desse total, 89% representam pessoas físicas, sendo que dessas, 44% possuem algum tipo de operação de crédito com a cooperativa, enquanto que, no segmento de pessoas jurídicas essa proporção salta para 67% (BACEN, 2015).

Relacionado ao número total de cooperativas no Brasil, observa-se (Figura 2) que sua ascensão foi expressiva a partir da década de 80, quando apresentava 430 cooperativas de crédito já em 1990 esse número era de 806, e na década de 90, saltando de 806 para 1330 no ano de 2010. Essa ascensão é atribuída principalmente ao fim da regulamentação, bastante restritiva, imposta na época pelo Governo Federal, ao cooperativismo.

Atribui de forma similar, a razão pela ampla expansão das cooperativas de crédito a partir da década de 90, justificando como o processo mais significativo para tal, Lopes (2001, p.19) diz, “com a promulgação da Constituição Federal de 1988, na

qual, entre outras coisas, retirou-se a exigência de autorização governamental para a constituição de cooperativas”.

Figura 2 - Número cooperativas de crédito



Fonte: elaborado pelo autor, 2018.

Atualmente o número de cooperativas de crédito, singulares, vem sofrendo uma redução. Ao fim de 2017, o número total de cooperativas singulares existentes foi de 986. Justifica-se essa diminuição pelo grande número de fusões e incorporações que se tem entre singulares, podendo ocorrer por diversos motivos, desde o ganho de escala até a má administração por parte de alguma delas, sendo englobadas por sistemas, centrais ou federativos, iguais ou diferentes.

Todavia, a redução no número de cooperativas singulares, não nos permite deduzir que há redução ou estagnação no crescimento das mesmas no sistema financeiro nacional, pelo contrário, há constante crescimento na participação no mercado financeiro, como descrito anteriormente. Conforme a Tabela 1, vemos que até o ano de 2014, houve crescimento expressivo das cooperativas de crédito. Do ano de 1994 até o ano de 2014, as cooperativas saltaram de ,0,7% para 3,5% do total de Patrimônio Líquido de todas as instituições financeiras; de 0,2% para 2,1% dos Ativos; de 0,1% para 3,1% do total de Depósitos; e de 0,3% para 2,9% das Operações de Crédito no país (BACEN, 2015).

Tabela 1 - Participação do cooperativismo de crédito no sistema financeiro nacional

Ano	Participação no patrimônio líquido	Participação nos ativos	Participação nos depósitos	Participação nas op. de crédito
1994	0,7%	0,2%	0,1%	0,3%
1995	1,4%	0,2%	0,1%	0,4%
1996	1,3%	0,3%	0,3%	0,5%
1997	1,6%	0,4%	0,5%	0,7%
1998	1,6%	0,5%	0,6%	0,9%
1999	1,8%	0,7%	0,8%	1,1%
2000	2,0%	0,8%	1,0%	1,2%
2001	2,0%	0,9%	1,3%	1,6%
2002	2,2%	1,0%	1,5%	1,8%
2003	2,2%	1,3%	1,8%	2,1%
2004	2,6%	1,4%	1,4%	2,3%
2005	2,9%	1,5%	1,4%	2,3%
2006	3,2%	1,9%	1,8%	2,8%
2007	2,6%	1,5%	1,4%	2,4%
2008	2,1%	1,3%	1,3%	2,6%
2009	2,4%	1,5%	1,4%	2,6%
2010	2,4%	1,6%	1,7%	2,4%
2011	2,6%	1,7%	1,8%	2,5%
2012	2,7%	1,8%	2,3%	2,3%
2013	3,1%	2,0%	2,6%	2,7%
2014	3,5%	2,1%	3,1%	2,9%

Fonte: BACEN, 2015.

Um fator corrente e de extrema importância, para crescimento do cooperativismo de crédito, é o estímulo do Governo Federal que tem tido uma visão de “combate” à concentração de renda e acesso ao crédito, assim como o Banco Central do Brasil que tem manifestado apoio ao setor cooperativista. Uma característica do cooperativismo de crédito que é determinante para essa visão, é o fato da pulverização de empréstimos. Praticamente 56% dos contratos liberados pelas cooperativas ficam na faixa de valores até R\$ 3 mil reais (BCB, 2006).

Outro fator relevante, é a disseminação para as regiões desassistidas pelas demais instituições financeiras, conforme citado anteriormente. Essa disseminação pode vir desde gestão, ou seja, estratégia das cooperativas já existentes, como também a percepção por parte da população, carente de tais serviços, que criam sua cooperativa ou buscam as cooperativas já existentes para atender a região e seus anseios e necessidades. Além disso, o cooperativismo de crédito tem o impacto

financeiro de apresentar juros médios mais baixos que outras instituições financeiras (ANIBAL; KOYAMA, 2011).

Apesar desse crescimento constante, motivado por diversos fatores, alguns expostos anteriormente, o cooperativismo de crédito no país tem um longo caminho a trilhar se observar a representatividade desse setor em outras regiões. Por exemplo, do total de intermediações financeiras realizadas na União Europeia já nos anos 2000, 15% já eram através das cooperativas de crédito (GONÇALVES; JACQUES, 2016).

## 2.2 CRÉDITO

Crédito é conceituado por Schrickel (1995, p. 42), como “todo ato de vontade ou disposição de alguém de destacar ou ceder, temporariamente parte do seu patrimônio a um terceiro, com expectativa de que essa parcela volte a sua posse integralmente, após decorrido o tempo estipulado”. Porém, refletindo empiricamente, o ato de pôr à disposição o patrimônio próprio, sendo pessoa física ou jurídica, leva-nos a presumir, em quase a totalidade desses atos, que haverá uma compensação ou recebimento de benefício próprio da parte receptora com o conessor. Como diz Securato (2002, p. 18), “apesar de existirem empréstimo a título gratuito, ou seja, não oneroso a aquele que recebe o bem, normalmente associa-se que qualquer transação de empréstimo tem um preço remuneratório, a ser pago pelo tomador”.

Indagando mais, quanto a definição de crédito, porém estritamente na perspectiva de instituições financeiras, podemos dizer que crédito é a concessão de um valor monetário para um agente solicitante, com comprometimento de devolução futura deste mesmo valor acrescido uma contrapartida para com instituição concessora do valor inicial. Usualmente, se não efetivamente em todas as conjunturas de concessão de crédito em instituições financeiras, a contrapartida é um valor definido entre as partes, chamado de juros, que significa um percentual tempestivo (diário, mensal, trimestral ou anual) sob o valor devido para a instituição a ser liquidado em uma data pré-definida, ou ainda amortizada tempestivamente até a liquidação do valor total deliberado entre o agente conessor e o agente receptor. Estrictamente sobre essa perspectiva, para instituições financeiras, Silva (2006, p. 39) diz:

Em um banco, que tem a intermediação financeira como sua principal atividade, o crédito consiste em colocar à disposição do cliente (tomador de recursos) certo valor sob a forma de empréstimo ou financiamento, mediante uma promessa de pagamento numa data futura. Na verdade, o banco está comprando uma promessa de pagamento, pagando ao tomador (vendedor) um determinado valor para, no futuro, receber um valor maior.

Já a origem do crédito é encontrada desde a antiguidade, ainda que de formas mais rudimentares. Seu primeiro relato foi na região do Egito, onde existiam regulamentos da época que citavam práticas similares a instrumentos creditícios. Porém, de forma mais robusta, o crédito é relatado na Roma e na Grécia antiga onde existiam os chamados cambistas, que realizavam trocas de diferentes moedas sempre incluindo um benefício próprio para consigo (CONANT, 1967).

Outro aspecto a ser abordado sobre o crédito é a sua relevância para o desenvolvimento da sociedade, desde a antiguidade até os dias contemporâneos. O crédito é um instrumento simplificador para a transação de bens e serviços nos modelos de sociedades capitalistas que prosperaram até hoje. Alguns autores chegam a afirmar que é instrumento indispensável para a dinâmica nessas sociedades. Desde o século vinte são realizados trabalhos relevantes, sob a perspectiva de que o desenvolvimento econômico está associado ao crédito. Dentre esses trabalhos, pode-se citar Goldsmith (1969) e McKinnon (1973) que buscavam corroborar essa ligação, de crédito com desenvolvimento econômico. Um estudo amplo e correlativo a estes, foi realizado por King e Levine (1993) que se utilizou de uma amostra de 80 países para afirmar que o desenvolvimento econômico desses é relacionado positivamente conforme o nível do desenvolvimento financeiro dos mesmos.

Esse instrumento, o crédito, tenciona a economia para vertentes, cujo essas estão correlatas justamente com o bem-estar econômico e progresso da sociedade, como a manutenção da demanda efetiva em compatibilidade com o pleno emprego da força de trabalho. Conforme Keynes (1936, p. 59), em seu livro, *The General Theory of Employment, Interest, and Money*, afirma:

É também verdade que a concessão do crédito bancário estabelecerá as seguintes tendências: (1) aumento da produção, (2) produto marginal aumente de valor em termos da unidade salarial (que em condições de redução do retorno deve necessariamente acompanhar o aumento da produção), e (3) aumento da unidade salarial em termos monetários (isso é frequente acompanhado de um melhor emprego); E essas tendências podem afetar a distribuição do rendimento real entre os diferentes grupos.

Essas tendências são explicadas devido a dinâmica que o modelo econômico capitalista possui. O crédito realiza, principalmente, o financiamento para o consumo de bens e serviços, de pessoas físicas, e o financiamento para o inovação e expansão de produção, de pessoas jurídicas. Assim, aumentando volume de vendas, abertura de novos mercados, satisfação de clientes, criação de novos consumidores, entre outros, o crédito acaba sendo um catalizador para expansão gradual da economia (BECKMAN, 1962)

### 2.3 RISCO DE CRÉDITO

Crédito possui uma infinidade de significados, alguns citados no tópico anterior. Entretanto, para elucidar de forma mais vasta este tópico, é abordado mais sobre os seus significados. Crédito vem do latim *creditu* que significa acreditar ou confiar. Também, no Novo Dicionário Aurélio possui uma série de significados, como:

1. Segurança de que alguma coisa é verdadeira; 2. Boa reputação; 3. Autoridade, influência, valia, importância; 4. Fé na solvabilidade; 5. Facilidade para obter dinheiro por empréstimo ou abrir contas em casas comerciais; 6. Facilidade de conseguir adiantamentos de dinheiro para fins comerciais, industriais, agrícolas, etc.; 7. Soma posta a disposição de alguém em um banco, numa casa de comércio, etc., mediante certas vantagens; 8. O que o negociante tem a haver; 9. O haver de uma conta; 10. Direito de receber o que se emprestou; 11. Quantia correspondente a esse direito; 12. Autorização para despesas dada por autoridades que estabelecem, votam ou regulamentam orçamentos; 13. Troca de bens presentes por bens futuros<sup>1</sup>.

Constata-se, a partir da origem da palavra crédito e de seus significados na norma linguística e lexicológica brasileira, que crédito está associado com confiança, segurança, reputação e outros sinônimos interligados. Em referência ao conceito de crédito, pode-se citar que esses – confiança, segurança, reputação e seus sinônimos – são componentes observados para a concessão do mesmo (SILVA, 2004). Ainda quanto essa relação, da concessão de crédito sob confiança, Keynes (1936, p. 101) diz:

---

<sup>1</sup> It is also true that the grant of the bank-credit will set up three tendencies - (1) for output to increase, (2) for the marginal product to rise in value in terms of the wage-unit (which in conditions of decreasing return must necessarily accompany an increase of output), and (3) for the wage-unit to rise in terms of money (since this is a frequent concomitant of better employment); and these tendencies may affect the distribution of real income between different groups. “(Tradução livre)”.

[...] também devemos ter em conta a outra face do estado de confiança, ou seja, a confiança das instituições de crédito em relação àqueles que procuram pedir-lhes empréstimos, por vezes descrito como o estado do crédito<sup>2</sup>.

Através dessa relação, da confiança do agente concesso com o agente receptor, pode-se afirmar que tal é uma condição fundamental para a consumação dos empréstimos pelas instituições financeiras. Por causa dessa relação, a concessão de crédito vem acompanhada do fator risco, tendo em vista que a confiança pode acabar vindo a ser violada. Incluindo esse fato na definição de crédito, Santos (2003) diz: “crédito refere-se à troca de um valor presente por uma promessa de reembolso futuro, não necessariamente certa, em virtude do fator risco”. Esse fator, que influencia nas tomadas de decisões dos mais variados assuntos, não unicamente na concessão de crédito, é definido por Gitman (1997, p. 33), “o risco pode ser definido como o grau de incerteza a respeito de um evento ou mesmo uma possível probabilidade de perda”. O risco é segmentado em diversos setores, porém irá se abordar somente o risco de crédito.

Risco de crédito é o risco envolvido no segmento de crédito. Seria um determinado nível, mensurável, de incerteza que há de não se obter de volta o patrimônio e o valor remuneratório deliberado entre as partes. Segundo Soares (2005, p. 14), risco de crédito é definido como:

[...] uma medida numérica da incerteza relacionada ao recebimento de um valor contratado e/ou compromisso, a ser pago por um tomador de empréstimo, contraparte de um contrato ou emissor de um título, descontadas as expectativas de recuperação e realização de garantias.

Compreendendo a importância do crédito para o desenvolvimento econômico de uma sociedade, como apontado anteriormente. A dimensão da relevância que é para a rentabilidade das instituições financeiras, tanto que essa indústria tem como um desafio tornar o crédito disponível na mais larga escala possível (LEWIS, 1992). E que a concessão de crédito está estritamente relacionada com exposição ao risco de crédito, notamos a dimensão do risco que qualquer instituição financeira está submetida. Essa maior exposição é ocasionada pelas inúmeras concessões de crédito.

---

<sup>2</sup> [...] thus we must also take account of the other facet of the state of confidence, namely, the confidence of the lending institutions towards those who seek to borrow from them, sometimes described as the state of credit “(Tradução livre)”.



Tem-se a ciência de que os não cumprimentos de pagamentos de empréstimos, a partir de um determinado nível, nos prazos e valores estipulado entre as partes, acarretaria em algum momento em dificuldade financeira para a instituição honrar seus compromissos com os aplicadores, podendo levar à sua insolvência. Conforme Saunders (2000, p. 108) diz sobre a insolvência de instituições financeiras, “ocorre sempre que os recursos próprios ou internos são insuficientes para cobrir perdas ocorridas em função dos demais riscos”.

Cabe citar, a respeito dos não cumprimentos de honrar as liquidações de concessões de crédito, que a crise de 1929, a maior crise da história até aquele momento, teve origem devido a esse tipo de ocorrência em grande escala. As instituições financeiras e o próprio governo americano realizaram volumosas somas de empréstimos, principalmente para países europeus, que se utilizaram desses recursos para o financiamento bélico da primeira guerra mundial e também para a reconstrução desses países assolados por tal. Posteriormente a isso, os agentes não possuíam condições financeiras para honrar com as amortizações e liquidações dessas concessões de crédito, devido à grande parte do volume emprestado ter sido utilizado para fins bélicos e não sendo produtivo para geração de rentabilidade futura, simplesmente um gasto sem retorno financeiro, e também, essas nações em reconstrução não conseguiam honrar os compromissos devido à falta de capacidade financeira em que se encontravam (PETIT; DE VEYRAC, 1945).

Os não cumprimentos de amortizações e liquidações em prazos estipulados, pune também os bons agentes captadores de recursos, no caso de instituições financeiras que não segmentam os clientes entre graus de riscos diferentes na concessão de crédito (ARAÚJO, 1996). Esse prejuízo causado aos bons captadores de recursos junto as instituições financeiras, é devido a inadimplência que impacta na taxa de juros de mercado, constantemente. Esta relação foi confirmada pelo estudo de Amaral, Oreiro e Silva (2012), onde questionavam por que a taxa de juros é tão alta no Brasil, e um dos fatores que tiveram relação positiva para o grande *spread* bancário no país foi a inadimplência.

Na década de 80, o Comitê de Supervisão Bancária do Banco de Compensações Internacionais preocupando-se com os riscos englobados no Sistema Financeiro Internacional, propôs algumas diretrizes para valores mínimos de capital e riscos de ativos, e foi dessas diretrizes que surgiu o primeiro Acordo de Basileia. Em 2004 novas orientações foram recomendadas, essas englobadas no chamado

segundo Acordo de Basileia, esse se sintetizava em três pilares. O primeiro pilar resumia novos aspectos do patrimônio líquido exigido com base em mensurações de risco de mercado, operacional e crédito. O segundo pilar abordava a regulamentação e fiscalização dos agentes sobre as instituições financeiras. Já o terceiro pilar era referente a divulgação e transparência dos participantes do mercado financeiro sobre seus riscos e sua gestão. Mais recentemente, 2010, sucedeu o terceiro Acordo de Basileia, onde no Brasil, iniciou-se seu cronograma de implementação a partir de 2013. Em linhas gerais, o último acordo traz novas medidas que atentam para o desincentivo da alavancagem financeira para os agentes do mercado financeiro (TAVARES, 2014).

Como forma de evitar-se a concessão de crédito de forma indiscriminada, ou seja, sem qualquer tentativa de discriminação entre bons e maus pagadores, por parte de instituições financeiras e seguindo recomendações do acordo Basileia, o Banco Central do Brasil por meio da Resolução, nº 1559, de 22 de dezembro de 1988, diz:

IX - É vedado às instituições financeiras:

- a) realizar operações que não atendam aos princípios de seletividade, garantia, liquidez e diversificação de riscos;
- b) renovar empréstimos com a incorporação de juros e encargos de transação anterior, ressalvados os casos de composição de créditos de difícil ou duvidosa liquidação;

Considerando-se a relevância, para as instituições financeiras, de se conceder créditos a bons pagadores mitigando ao máximo o risco de crédito, a ameaça de continuidade da instituição, como também, o cumprimento da regulamentação que veda a empresa em realizar concessões de créditos que não atendem aos princípios de seletividade, garantia, liquidez e diversificação de riscos, essas instituições buscam maneiras de se mensurar o risco dos tomadores para a decisão de se aprovar ou não as operações solicitadas. Essa atividade, a análise de solicitação de crédito, de uma empresa busca determinar se deve ser concedido crédito a um cliente, e quais os limites monetários devem ser impostos, ou não. A partir disso, quais seriam as informações necessárias e suas fontes para que se possa avaliar a capacidade creditícia de um cliente (GITMAN, 2003).

Para Scherr (1989) o modelo mais tradicional de uma organização de informações sobre a possibilidade de pagamento de um cliente é caracterizado pelas suas cinco dimensões, sendo conhecido como os cinco "C" do crédito":

**Caráter:** Representado pelo histórico de pagamento dos compromissos financeiros e contratuais. Para Silva (1988) o caráter está associado com a intenção do cliente de pagar sua dívida, ou seja, é um conjunto de qualidades do devedor face aos seus hábitos de pagamento. Faz parte desse conjunto pontualidade nos compromissos anteriores, a tradição de pagamento e as restrições existentes sobre seu nome.

**Capacidade:** Representada pelo potencial financeiro para honrar os compromissos. Normalmente, são analisados dados relativos a renda e aplicações financeiras para as pessoas físicas, e índices de liquidez ou endividamento para as pessoas jurídicas.

**Capital:** Representado pela solidez patrimonial ou saúde econômico-financeira do devedor. Nesse item, em geral, são avaliados patrimônio pessoal, nas pessoas físicas, e balanço patrimonial, no caso de pessoas jurídicas. No balanço, normalmente é observado o índice de endividamento ou liquidez. Contudo, é importante ressaltar que esses índices representam informações passadas.

**Colateral:** Representado pelas garantias adicionais oferecidos pelo devedor para atendimento de demandas de bancos. Devem ser aplicados nos colaterais os mesmos procedimentos de avaliação de crédito utilizada no solicitante. Em geral as garantias ultrapassam o próprio valor da dívida para proteger os credores contra problemas associados à liquidez dos ativos garantidores.

**Condição:** Representada pelos fatores econômicos vigentes e as características dos indivíduos que podem aumentar ou diminuir a capacidade de pagamento do devedor. Para as pessoas jurídicas esses fatores estão normalmente associados às políticas governamentais, conjunturas nacional e internacional, concorrências, etc. Apesar dos fatores não serem controlados pelos devedores, eles afetam de maneira importante suas capacidades de honrar os compromissos.

Esse modelo tradicional, de organização de informações sobre a possibilidade de pagamento de um cliente, conhecido como "5C's do crédito", é referência até os dias contemporâneos, e foi abordado por Weston e Brigham em 1972, na obra *Managerial Finance*.

Novos autores acrescentaram uma nova dimensão, até pela usabilidade empírica como ferramenta de análise em setores de créditos de instituições financeiras, transformando-o em os seis "C's" do crédito". Essa sexta dimensão é

conhecida como Conglomerado, e o incremento desta nova dimensão é derivado da má fé por parte de alguns agentes, solicitantes de crédito. Esses se utilizam de mecanismos de repasse por parte de conglomerados de empresas coligadas, ou por um grupo de pessoas relacionadas, como por exemplos familiares, ou o enlevar de pessoa física com jurídica. Esses agentes solicitantes apresentam boa ou razoável condição financeira, e posteriormente a liberação do crédito realiza o repasse do valor para outra empresa ou pessoa, cujo se encontra em deterioração financeira. Este mecanismo de má fé, se não analisado pelo setor de crédito, pode causar equívoco na análise do risco do crédito, tendo em vista a falta de veracidade que o risco dessa concessão de crédito representa realmente para a instituição financeira (SILVA, 2000).

Outro ponto atentado, pelas instituições financeiras e empresas em geral, referente ao risco de crédito, é a forma de gerir esse risco. Essa gestão do risco de crédito é estabilidade em três pontos. O risco cliente cujo é a probabilidade de *default* por parte do tomador, esse risco que é analisado nos “C’s” de crédito. O risco da operação, cujo a perda deriva da composição do crédito, como garantias, condições e produto. E o risco de concentração, que é a perda decorrente da composição e características da carteira, como exemplo setor de atividade dos clientes, produto bancário ou região geográfica (TAVARES, 2014).

## 2.4 CREDIT SCORING

Um dos aperfeiçoamentos de análise dos C’s do crédito foi a criação do *rating*, que consistem em estabelecer um determinado nível para o risco de crédito de uma pessoa, ou empresa, baseado em variáveis relacionadas a Caráter; Capacidade; Capital; Colateral; Condição. Conforme Blatt (1999) diz:

[...] um *rating* de crédito é um sistema de classificação de risco creditício, realizada muitas vezes através de letras e/ou números. O *rating* é um exclusivo indicador que possibilita visualizar a capacidade financeira de uma empresa e o nível de risco associado.

Esse sistema de classificação é empregado principalmente nas instituições financeiras e agências classificadoras de *rating*. Nas instituições financeiras é utilizado para avaliação do risco de crédito, assim como a tomada de decisão da aprovação ou

não, de tomadores de crédito com as instituições. Esse risco avaliado tem por expectativa mensurar da melhor forma possível a chance de um iminente *default* por parte desses tomadores (ASSAF; BRITO; CORRAR, 2009).

Ainda, referente a classificação de *rating* por parte das instituições financeiras, estas são obrigadas a efetuarem a classificação em operações concedidas e em andamento (mesmo após a concessão deve-se continuamente efetuar a classificação de *rating*), para que se mensure e atualize o provisionamento de possíveis não recebimentos que as instituições podem vir a sofrer. Segundo o Banco Central do Brasil, por meio da Resolução, nº 2682, de 21 de dezembro de 1999, o provisionamento deve ser feito nos percentuais e *ratings*, conforme a Quadro 1.

Quadro 1 - *Rating*\percentual de provisionamento

<b>Rating</b>	<b>Percentual Provisionado</b>
AA	0%
A	0,5%
B	1%
C	3%
D	10%
E	30%
F	50%
G	70%
H	100%

Fonte: elaborado pelo autor, 2018.

Esse acompanhamento mensal e alteração de classificação de *rating*, consigo o seu respectivo provisionamento, deve ser empregado, segundo a mesma Resolução nº 2682, de 21 de dezembro de 1999, com nível de classificação mínima baseado nos dias em atraso que as operações possam estar vindo a ocorrer, e segue as seguintes regras:

- a) atraso entre 15 e 30 dias: risco nível B, no mínimo;
- b) atraso entre 31 e 60 dias: risco nível C, no mínimo;
- c) atraso entre 61 e 90 dias: risco nível D, no mínimo;
- d) atraso entre 91 e 120 dias: risco nível E, no mínimo;
- e) atraso entre 121 e 150 dias: risco nível F, no mínimo;
- f) atraso entre 151 e 180 dias: risco nível G, no mínimo;
- g) atraso superior a 180 dias: risco nível H.

Um outro método empregado para liberação de crédito, a pessoas jurídicas e físicas, em instituições financeiras é o *credit score*. O método é de algum modo similar

a classificação de *rating*, porém no *credit score* utiliza-se de instrumentos estatísticos para se conceder pesos as variáveis utilizadas no modelo, e ao fim atribui-se um valor proporcional ao seu risco de *default*. Segundo Vicente (2001, p. 49):

*Credit Score* pode ser definido como o processo de atribuição de pontos às variáveis de crédito mediante aplicação de técnicas estatísticas. Trata-se de processo que define a probabilidade de que um cliente com certas características, considerada desejáveis ficando a critério da instituição dar alçada operacional ou não para gestor atribuir o crédito. Essa técnica estabelece uma regra de discriminação de um determinado cliente solicitante de crédito.

O modelo estatístico do *credit score* tem como característica ponderar pesos em variáveis selecionadas, essas que são de suma importância para acurácia do modelo. Entretanto seria impossível se dispor da totalidade de variáveis possíveis em uma concessão de crédito, assim, apesar de o modelo ser utilizado em larga escala até os dias de hoje através de diversas técnicas estatísticas e trazer diversos benefícios, é utópico pensar que qualquer instituição teria um modelo com 100% de acurácia. Chaia (2003, p. 20) diz:

Sua ideia básica é identificar, através de técnicas estatísticas, os principais fatores que determinam a probabilidade de inadimplência. Apesar do *credit scoring* ser um processo matemático, ele não elimina a possibilidade de se recusar um mau pagador ou de se aceitar um mau pagador.

Para o desenvolvimento do modelo de *credit score* é necessário a segregação e desenvolvimento de etapas. Estas que são: planejamento e definições; identificação de variáveis; planejamento da amostra; determinação do *score* (método estatístico); validação e verificação de performasse do modelo; determinação do ponto de corte (SICSU, 1999), apresentadas na Figura 3 abaixo:

Figura 3 - Etapas do desenvolvimento do *credit score*



Fonte: elaborado pelo autor, 2018.

Planejamento e definições remete-se a etapa onde se examinará às possíveis diferenciações de modelos, visão geral da população, disponibilidade de dados e, principalmente, a definição de inadimplente para o modelo de *credit score* a ser estimado. Identificação de variáveis, seria quais os dados disponíveis a serem utilizados na estimação do modelo. Planejamento da amostra, é a definição da amostra a ser empregada, dentre a sua população total. Método estatístico a ser utilizado, tendo em vista a grande variedade e diferenças de métodos possíveis a serem empregados. Verificação de performance, diz respeito a testes de medidas de desempenho do modelo final estimado. E a determinação do ponto de corte do *credit score* para a classificação de inadimplência do modelo, questão subjetiva a política e as utilidades que a empresa venha a necessitar com a utilização do modelo.

Dentre as técnicas estatísticas utilizadas para o desenvolvimento dos modelos, pode-se citar a aplicação de uma vasta variedade, em trabalhos recentes na análise de risco de crédito no Brasil. A regressão múltipla (SANTOS; FAMÁ, 2007); Método multicritério AHP (GARTER; MOREIRA; GALVES, 2008); Redes neurais (STEINER et al. 2007); Regressão Logística (ASSAF; BRITO; CORRAR 2009); Análise discriminante (GUIMARÃES; MOREIRA, 2008) entre outros.

Silva (2014) em seu trabalho expos, através de uma revisão de literatura, que à técnica mais utilizada, para o desenvolvimento de estudos no âmbito acadêmico, nacionais e internacionais, foi a regressão logística. Apresentou também, que a grande diferenciação dos estudos, nacionais para os internacionais, é que os trabalhos nacionais estão focados no fornecimento de modelos voltados para a seleção de variáveis, e posteriormente a mensuração da acurácia. Enquanto que os trabalhos internacionais tendem a realizar diversos modelos, com vários conjuntos de dados e situação, tentando apresentar qual é o modelo mais eficiente.

Guimarães e Chaves Neto (2002) propuseram, para a análise e concessão de crédito, um método através da regressão logística e função linear discriminante. O estudo estava restrito a concessão de limites de crédito para cartões de crédito de pessoas físicas, em uma administradora de cartões. Os resultados demonstraram que o modelo logístico obteve maior acurácia, em relação ao modelo discriminante.

Um outro trabalho, de Louzada Neto, Ferreira e Diniz (2012), estudou em um banco brasileiro, os impactos - acurácia, sensibilidade e especificidade - de amostras desproporcionais na estimação de modelos de concessão de crédito sob o método de

regressão logística. Foi sugerido que essa desproporcionalidade não causa prejuízo, sob a acurácia, sensibilidade e especificidade.

Dentre as variáveis utilizadas para o desenvolvimento dos modelos de análise de risco de crédito, restringindo-se a perspectiva de modelagem para pessoas físicas, constataram-se 39 variáveis e suas respectivas frequências, em 11 trabalhos analisados, conforme o Quadro 2.

Quadro 2 - Variáveis e frequências

Variável	Frequência
Renda Mensal	9
Estado Civil	8
Idade	8
Profissão	8
Tempo de Vínculo no Emprego Atual	8
Patrimônio (Bens Móveis e Imóveis)	6
Sexo	6
Tipo de Linha de Crédito \ Finalidade do Crédito	6
Garantias e Avalistas	5
Percentual de Comprometimento da Renda com Dívidas	5
Número de Dependentes	5
Número de Parcelas \ Período de Contrato	5
Tempo na Residência Atual	5
Idoneidade do Cliente Pessoa Física (CCF, Serasa, Outros)	5
Aplicações Financeiras na Instituição	4
Tempo de Relacionamento com a Instituição Financeira	4
Endividamento no Sistema Financeiro Nacional (BACEN)	3
Plano de Saúde Familiar	3
Primeiro Empréstimo \ Volume de Operações Creditícias	3
Valor do Empréstimo	3
Comportamento em Créditos Anteriores	3
Comportamento da Conta Corrente	2
Grau de Instrução	2
Localidade da Residência	2
Número de Cartões de Crédito na Instituição Financeira	2
Plano de Seguro para Veículos	2
Plano de Seguro Residencial	2
Posse de Telefone Residencial	2
Taxa de Juros da Operação	2
Valor da Parcela	2
Brasileiro\Estrangeiro	1
Natureza Contratual da Empresa do Funcionário	1
Indicação de Cliente	1
Grau de Rotatividade dos Produtos em Épocas de Recessão	1
Perspectiva de Relacionamento Futuro	1
Posse de Fone Comercial	1
Renda Informal	1
Renda Mensal Líquida do Conjugue	1
Tempo de Cartão de Crédito	1

Fonte: elaborado pelo autor, 2018.



### 3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa se utilizou dos referenciais teóricos de livros, dissertações, teses, sites, artigos científicos, leis e regulamentações, assim como, das políticas, instrumentos, dados e experiências empíricas da cooperativa estudada, na análise e concessão de crédito aos cooperados pessoas físicas.

O trabalho busca estimar um modelo de *credit score* para a análise e liberação automática de crédito a pessoas físicas, por meio de modelagem estatística, cujo será obtido através das informações cadastrais, de utilização e de restrições de mercado. O estudo estará delimitado para o uso empírico, exclusivamente, na cooperativa de crédito estudada, além de contribuição para a literatura nacional do tema, no meio cooperativista de crédito.

#### 3.1 ESPECIFICAÇÃO DO PROBLEMA

Diante do contexto apresentado no problema de pesquisa e do referencial teórico, busca-se responder o seguinte problema:

**Como estimar um modelo de *credit score*, para solicitantes de crédito pessoas físicas, na cooperativa de crédito estudada, que acarrete em um aumento de automação na análise e concessão de crédito?**

#### 3.2 CARACTERIZAÇÃO DA PESQUISA

A natureza da pesquisa é aplicada, uma vez que objetiva estimar um modelo empírico de *credit score*, para a análise e liberação automática de crédito a solicitantes de uma cooperativa de crédito.

Quanto à abordagem da pesquisa, será quantitativa, tendo em vista a utilização de variáveis quantitativas para se mensurar o grau de confiabilidade dos solicitantes de crédito, através de um modelo de *credit score*, e posteriormente classificá-los em futuros bons ou maus pagadores. Conforme Fonseca (2002, p. 20): “Diferentemente da pesquisa qualitativa, os resultados da pesquisa quantitativa podem ser quantificados”, e completa: “[...] os resultados são tomados como se constituíssem um

retrato real de toda a população alvo da pesquisa. A pesquisa quantitativa se centra na objetividade”.

O objetivo da pesquisa é de caráter explicativo, pois tenta estabelecer dentre os solicitantes de crédito, quais serão adimplentes e inadimplentes, por meio das variáveis dispostas na pesquisa.

Quanto aos procedimentos utilizados foi utilizada a pesquisa bibliográfica. Bibliográfica devido ao levantamento e análises dos livros, dissertações, teses, sites e artigos científicos utilizados. A pesquisa bibliográfica é ponto de partida de qualquer trabalho científico (FONSECA, 2002).

### 3.3 COLETA E TRATAMENTO DE DADOS

Esta seção trará as informações referentes a base de dados utilizada na pesquisa. Irá apresentar o período em análise, os produtos de crédito, o número de propostas aprovadas pela cooperativa e as categorias que foram utilizadas. Além disso, irá discutir a definição da variável resposta a ser utilizada no modelo de *credit score*.

#### 3.3.1 Dados e produtos de crédito

Foram selecionadas todas as solicitações de crédito na cooperativa em estudo, do período de 2013 até 2015. As solicitações do período, cujo foram aprovadas, já foi transcorrido o período mínimo de 18 meses de sua aprovação, referente as últimas solicitações de crédito aprovadas em dezembro de 2015, e o período máximo de 54 meses, referentes as primeiras aprovações em janeiro de 2013, assim sendo é possível avaliar o comportamento de inadimplência dentro de um espaço de tempo independente de sua liquidação ou não.

Foi verificado um total de 6130 solicitações de crédito aprovadas para pessoas físicas. Essas solicitações são compostas por uma diversidade total de 62 linhas de crédito diferentes, e foram segmentadas em 5 categorias: Financiamentos de bens com garantias dos mesmos; Crédito com alguma garantia monetária; Créditos sem garantias reais; Crédito com garantia de bens; Antecipação de garantias futuras. As solicitações aprovadas e o número por categoria constam na Tabela 2, onde é

possível, também, verificar a quantidade de propostas aprovadas que obtiveram inadimplência acima de 14, 30 e 90 dias.

Tabela 2 - Propostas, categorias e inadimplência 1

2013-2015	Aprovada	>14	>30	>90
Financiamentos de bens com garantias dos mesmo	1149	174	96	36
Crédito com alguma garantia monetária	1762	74	11	2
Créditos sem garantias reais	2467	334	169	56
Crédito com garantia de bens	25	9	8	4
Antecipação de garantias futuras	729	59	19	6
<b>Total</b>	<b>6130</b>	<b>650</b>	<b>303</b>	<b>104</b>

Fonte: elaborado pelo autor, 2018.

A categoria crédito com garantia de bens, acabou sendo excluída da pesquisa devido ao baixo volume de solicitações, o que traria prejuízo a modelagem. Já a categoria crédito com alguma garantia monetária, foi excluída pelo motivo de baixíssima inadimplência, o que acabaria por não trazer resultados estatísticos significativos para a modelagem do *credit score*.

Posteriormente a retirada das duas categorias, obtivemos um total de 4345 solicitações que foram aprovadas. Na Tabela 3 é possível visualizar a distribuição das propostas pelas três categorias restantes e o número de inadimplentes, de acordo com as variedades de inadimplência, 14, 30 e 90 dias.

Tabela 3 - Propostas, categorias e inadimplência 2

2013-2015	Aprovada	>14	>30	>90
Financiamentos de bens com garantias dos mesmo	1149	174	96	36
Antecipação de garantias futuras	729	59	19	6
Créditos sem garantias reais	2467	334	169	56
<b>Total</b>	<b>4345</b>	<b>567</b>	<b>284</b>	<b>98</b>

Fonte: elaborado pelo autor, 2018.

A categoria de financiamento de bens com garantia dos mesmos é composta por 16 linhas de crédito, cujo são os financiamentos de veículos, imobiliários e equipamentos profissionais. A categoria antecipação de garantias futuras é composta por 7 linhas de crédito, a antecipação de convênios e recebíveis debitados na cooperativa, a antecipação de imposto de renda a receber na cooperativa e o crédito consignado. A categoria de créditos sem garantias reais é a mais ampla dentre as três categorias. Essa categoria foi nomeada de crédito sem garantias reais motivada pela

essência das linhas, cujo é diferente das outras duas categorias, que obrigatoriamente possuem uma garantia para a cooperativa na concessão de crédito, essa não tem tal obrigatoriedade. Porém, nada impede que essas linhas de crédito venham a possuir garantias para a cooperativa. A categoria é formada pelas 23 linhas de crédito, cujo são aquisição de cotas em outras cooperativas de ramos diferentes, crédito destinado a congressos e treinamento profissional, ação sustentável, crédito pessoal, móveis, reforma residencial e comercial, viagem, IPTU, IPVA e *hot money*. No Quadro 3 abaixo é possível verificar mais detalhadamente a nomenclatura das linhas por cada uma das três categorias.

Quadro 3 - Linhas de crédito por categorias

<b>Financiamentos de bens com garantias dos mesmo</b>	<b>Créditos sem garantias reais</b>
Crédito Imobiliário-alienação - Jr-Pré+Corr Ajs	Aquisição de Cotas - Jr.-Pré + Corr. Ajustada
Crédito Imobiliário-alienação - Jr-Pré+Corr. Ajs PF	Aquisição de Cotas - Jr-Pré+Cor Ajs - CDI
Crédito Imobiliário-alienação - lib comp - JrPré+Ajs PF	Aquisição de Cotas Partes - Price
Crédito Imobiliário-alienação - lib comp - JrPr+Ajs	Congressos - Price
Crédito Imobiliário-Jr. Pré + Corr. Ajustada - INPC	Crédito Ações Sustentáveis - Price
Crédito Imobiliário-alienação - SAC + Corr. Mon.	Crédito Pessoal -- Folha de Pgto - Price
Crédito Veículos Híbridos e Adaptados - Price	Crédito Pessoal - Jr.-Pré + Corr. Ajustado - CDI
Motocicleta Nova - Price	Crédito Pessoal - parcela única - Jrs. Pré
Veículo Próprio - Price	Crédito Pessoal - Price
Veículos Novos - Price	Crédito Pessoal - SAC
Veículos Novos - SAC	Crédito Pessoal Aniversário c/Aval - Price
Veículos Usados - Jr.-Pré + Corr. Ajus. - CDI	Crédito Pessoal Aniversário s/Aval - Price
Veículos Usados - Price	Crédito Pessoal Online - Price
Equip de Trab Profissional - Jr-Pré + Corr Ajs - CDI	Cursos de Aperfeiçoamento/Especialização - Price
Equip. de Trabalho Profissional - Price	Eleto/Eletrônicos e Móveis - Price
Equip. de Trabalho Profissional - SAC	Hot Money - Jr.-Pré
<b>Antecipação de garantias futuras</b>	<b>I.R. a pagar - Price</b>
Antecipação de Produção-Jr Pré + Cor Ajus - CDI	Iptu PF - Price
Antecipação de Produção - Price PF	Ipva - Price
Antecipação de Produção - Rotativo PF	Reforma Res. e Comerc. - Jr-Pré+Corr Ajustada - CDI
Antecipação de Produção - SAC PF	Reforma Residencial e Comercial - Price
Antecipação de Recebíveis - Rotativo - Cielo PF	Reforma Residencial e Comercial - SAC
Crédito Consignado - Price	Viagem e Lazer - Price
I.R. a restituir - Price	

Fonte: elaborado pelo autor, 2018.

### 3.3.2 Variável resposta

A variável resposta corresponde em saber se a pessoa física, que teve sua solicitação de crédito atendida, irá no decorrer do período de pagamento incorrer em inadimplência ou não. Isso nos remete a definir o que seria uma pessoa inadimplente na cooperativa de crédito estudada. Na literatura tem-se uma vasta diversidade a respeito desta definição, mas em um contexto geral a definição de inadimplência baseia-se no conceito de não cumprir com uma obrigação com a instituição financeira. Próximo disto, Westgaard e Wijst (2001, p. 339) dizem: "[...] entrar em *default* é fracassar em pagar uma quantia devida a um banco".

Alguns outros autores definem mais precisamente a inadimplência, inserindo um prazo máximo de dias de atraso, antes que você seja considerado um inadimplente, ou então decidem por uma definição já utilizada no mercado de crédito em geral. Andrade (2007, p. 1573) cita a inadimplência em nosso mercado nacional a inclusão no cadastro pessoa física (CPF) no sistema Serasa, cujo: "A definição de default é qualquer registro negativo disponível publicamente ou nos arquivos privados do Serasa"<sup>3</sup>.

Mais abrangente ainda é a definição do BCBS (*Basel Committee on Banking Supervision*) (2006, p. 100):

Considera-se ter ocorrido default em relação a um devedor específico quando um ou ambos os eventos seguintes tenham acontecido:

- O banco considera improvável que o devedor pague na totalidade suas obrigações ao conglomerado financeiro sem que este tenha que recorrer a ações tais como a realização de garantias (se possuir);
- O devedor está atrasado em mais de 90 dias em alguma obrigação material com o conglomerado financeiro. Saques a descoberto são considerados como operações em atraso quando o cliente infringir um limite recomendado ou tenha lhe sido recomendado um limite menor que a dívida atual.

Uma outra questão quanto a definição de inadimplência para variável resposta do modelo de *credit score* é quanto à política de crédito que cada instituição utiliza, algumas adotam critérios mais austeros, tendo um perfil mais comedido, já outras adotam critérios menos austeros e abrandam a definição.

---

<sup>3</sup> The definition of defaults is any negative report registered in the publicly available files and the private files managed by Serasa (Tradução Livre).

Para o trabalho em questão, foi rodado dois modelos com variáveis respostas de definições diferentes. Uma baseia-se em utilizar a definição de que inadimplentes foram os solicitantes que incorreram em um período superior a 30 dias de atraso, assim sendo, utilizando-se de uma definição mais austera. Já o outro modelo irá possuir uma variável resposta com solicitantes que incorreram em um período superior a 90 dias de atraso.

A variável respostas foi modificada para ser uma variável *dummy*, como a própria regressão logística necessita, e no modelo mais brando assumiram os valores 1 os solicitantes que incorreram no período de atraso maior que 30 dias durante o decorrer dos pagamentos. Já no segundo modelo, o mais brando, assumiram os valores 1 os solicitantes que incorreram em um período de atraso maior que 90 dias.

### **3.3.3 Variáveis explicativas**

As variáveis explicativas são compostas por informações do perfil (ou cadastro) do cooperado, do comportamento (ou utilização) e de restrições de mercado. Essas variáveis serão utilizadas nas modelagens para se verificar a significância e peso sobre a variável resposta, ou seja, quais dessas informações serão relevantes para determinar a inadimplência, de acordo com o modelo estimado, de 30 ou 90 dias, dos solicitantes de crédito na cooperativa. Os modelos se utilizaram das mesmas variáveis independentes.

As variáveis com informações de cadastro são compostas pelas variáveis: Conta individual ou conjunta; Idade do solicitante; Número de anos como cooperado (tempo de relacionamento); Sexo do solicitante; e Valor da renda.

As variáveis com informações de utilização são: Categoria do empréstimo (3 categorias); Média de saldo em conta corrente nos últimos 3, 6 e 12 meses; Cheques devolvidos nos últimos 12 meses motivados por insuficiência de saldo; Número de parcelas do empréstimo solicitado; Valor do limite de cheque especial na cooperativa; Bens como garantias na solicitação de empréstimo; Valor em aplicações na cooperativa; Valor em cotas na cooperativa; Garantias monetárias na solicitação do empréstimo; e o valor total solicitado do empréstimo.

Quanto as informações de restrição de mercado, compreende-se somente o Valor Serasa, que o solicitante venha a possuir. O modelo apresenta então um total

de vinte variáveis explicativas e a variável resposta. Pode-se verificar no Quadro 4, onde estão discriminadas as variáveis por categoria de informação.

As variáveis explicativas Conta (individual ou conjunta), Categoria do Empréstimo, Sexo, Bens em Garantia e Valor Monetário em Garantia por se tratarem de variáveis qualitativas foram codificadas em variáveis *dummies* (0 e 1). Para cada uma dessas variáveis o número de categorias ( $k$ ), determinou o número de *dummies* ( $k - 1$ ) a ser empregado na regressão logística.

Quadro 4 - Variáveis

Variáveis	Categoria da Informação
Consulta Serasa do Solicitante	Restrição de Mercado
Conta Individual ou Conjunta	Cadastro
Idade	Cadastro
Categoria de Empréstimo /Linha de Crédito	Utilização
Média em Conta Corrente nos Últimos 3 Meses	Utilização
Média em Conta Corrente nos Últimos 6 Meses	Utilização
Média em Conta Corrente nos Últimos 12 Meses	Utilização
Número Anos como cooperado	Cadastro
Cheques Devolvidos nos Últimos 12 Meses	Utilização
Número de Parcelas Mensais do Empréstimo	Utilização
Número Máximo de Dias em Inadimplência	Variável Resposta
Saldo Devedor de Empréstimos na Cooperativa	Utilização
Sexo	Cadastro
Valor da Renda	Cadastro
Valor de Cheque Especial	Utilização
Bens como Garantia	Utilização
Valor em Aplicações	Utilização
Valor em Cotas	Utilização
Valor Monetário como Garantia	Utilização
Valor Solicitado	Utilização

Fonte: elaborado pelo autor, 2017.

Observa-se que o modelo poderia abranger uma maior quantidade de variáveis. Essa ausência, de maior número de variáveis, é motivada por alguns fatores: a falta de fidedignidade quanto ao preenchimento cadastral de determinadas variáveis, inexistência de informações não solicitadas pela cooperativa aos cooperados, problemas com busca de informações retroativas no software da cooperativa, impasses quanto a disponibilização de determinadas informações pela cooperativa e sigilos contratuais para repasse a terceiros.

### 3.4 SOFTWARES UTILIZADOS

Para uma visão geral, sumarização e produção do arquivo apto para utilização no software estatístico foi utilizado o *MS EXCEL 2016*, devido a sua ampla e suficiente funcionalidade para as demandas necessárias. Para os procedimentos estatísticos na estimação dos modelos foi utilizado o *software STATA (Data Analysis and Statistical Software)*, por atender todas as demandas estatísticas que a pesquisa vinha a recorrer.

### 3.5 MÉTODO DE ANÁLISE

Posteriormente as questões relativas a coleta e tratamento da base de dados, vamos à apresentação e discussões referentes a modelagem estatística para a estimação dos modelos de *credit score*.

#### 3.5.1 Regressão logística

A regressão logística diferentemente de outras regressões, cuja variável dependente manifesta-se na forma quantitativa, tem por peculiaridade estudar a ocorrência ou não ocorrência de um evento (FÁVERO, 2015). Logo, é representada na forma dicotômica, ou seja, assumindo valores de 1 e 0, dependendo da ocorrência ou não ocorrência do evento (GUIMARÃES; CHAVES NETO, 2002).

Na regressão logística ou análise *logit*, às variáveis explicativas e seus respectivos comportamentos é que descreverão a ocorrência ou não ocorrência do evento. As variáveis explicativas podem ser tanto variáveis métricas quanto variáveis *dummies*, e são definidas pelo pesquisador na coleta e tratamento dos dados.

O modelo da regressão logística pode ser representado pela função:

$$P(Y = 1) = \frac{1}{1 + e^{-(x)}}$$



Onde:

$P (Y=1)$  = Probabilidade relativa que um evento venha a ocorrer;

$e$  = número neperiano de valor aproximado de 2,718;

$$x = B_0 + B_{1i} * X_{1i} + B_{2i} * X_{2i} + \dots + B_{ni} * X_{ni}$$

Em que:

$B_0$  é a constante e  $B_1, B_2 \dots B_n$  serão os coeficientes estimados a partir das variáveis explicativas ( $X_1, X_2 \dots X_n$ ).

Logo, a regressão logística em si busca demonstrar a probabilidade de um evento ocorrer ou não, para isso é necessário a estimação dos coeficientes das variáveis explicativas.

Os coeficientes são calculados pelo método da máxima verossimilhança, cujo é o método mais utilizado para a regressão logística (SHARMA, 1996), e que novamente difere a regressão logística das tradicionais regressões, cujo se utilizam, em sua maioria, do método de mínimos quadrados para a estimativa dos parâmetros dos coeficientes. O método da máxima verossimilhança busca maximizar os parâmetros combinados com a probabilidade da amostra observada (HOSMER; LEMESHOW, 2004).

### 3.5.2 Avaliações dos modelos

Um dos motivos de a regressão logística ser tão utilizada em diversas áreas de pesquisa é devida a ausência da obrigatoriedade de uma série de pressupostos, que acabam incorrendo em outros métodos de análise, por exemplo às que se utilizam dos mínimos quadrados para estimação de coeficientes e consigo demandam distribuições t, distribuições F, R-quadrado, etc. (CORRAR et al., 2014).

Apesar dessa baixa exigência de pressupostos para a regressão logística, existe um ponto essencial, a necessidade de ausência de multicolinearidade entre as variáveis explicativas. Então, para a regressão logística é necessária a verificação e constatação da baixa colinearidade das variáveis independentes (HAIR JUNIOR et al., 2009).

Para a verificação de multicolinearidade pode ser realizado uma diversidade de testes, dois desses são o *Variance Inflation Factor* e a matriz de correlação. Com eles

é possível verificar se, e o quanto, às variáveis estão causando prejuízo na estimativa dos parâmetros, e assim, também, na regressão logística.

Como o trabalho busca gerar o melhor resultado para a predição de adimplência e inadimplência, torna-se necessário a verificação de quais variáveis explicativas são estatisticamente significantes para os modelos. O método *forward stepwise* foi o escolhido para se verificar e selecionar quais variáveis explicativas são significantes para os modelos. No método *forward stepwise* ocorrem simulações sequenciais de regressões, iniciando somente com a constante, adicionando às variáveis explicativas uma a uma, desde que dentro do nível de significância determinado, até que se encontre o modelo final que tenha maior qualidade. Para o trabalho foi definido um nível de significância estatística menor ou igual a 0,05.

Para a comparações de modelos pode-se citar três medidas. O Pseudo  $R^2$ , que é uma medida similar ao  $R^2$  utilizado em outros tipos de regressões, o valor varia de 0 a 1, e quanto mais próximo de 1 melhor será o ajuste do modelo. O Cox & Snell e o Nagelkerke também são medidas semelhantes, entretanto no Cox & Snell a medida acaba se limitando, tendo em vista que o valor nunca atingirá o valor máximo de 1. Nagelkerke alterou isto de modo que os valores pudessem estar entre 0 e 1 (FÁVERO, 2015).

Na verificação da qualidade do modelo final, pode-se utilizar do teste de Hosmer e Lemeshow. Esse teste é baseado na divisão da amostra em 10 partes e posteriormente, com um teste qui-quadrado, faz averiguar se houve diferenças significativas entre às classificações do modelo e às ocorrências reais da amostra para cada uma das 10 partes (CORRAR et al., 2014).

Outro método amplamente utilizado para se avaliar a acurácia do modelo é a chamada curva ROC (*Receiver Operating Characteristic*). Através do gráfico gerado, cujo é elaborado a partir da especificidade e a sensibilidade do modelo, é possível obter a AUC (*Area Under the Curve*) que representa a probabilidade de 2 casos, um evento e um não evento, de o evento ser classificado com um escore maior que o caso não evento. Logo, quanto mais se aproximar de 1 melhor a acurácia do modelo sob os dados utilizados (FLACH, 2010).

Por fim, uma outra ferramenta para a avaliação do modelo gerado é a matriz de classificação, diferente de outros métodos que se baseiam na significância estatística em si, a matriz de classificação traz uma perspectiva prática. Nela é possível verificar a precisão de classificação do modelo gerado, podendo ser utilizada

tanto para precisão da amostra utilizada quanto em amostras de validações. A matriz de classificação demonstra a quantidade e o percentual de classificações corretas, tanto para os casos de ocorrência do evento como não ocorrência (FÁVERO, 2015).

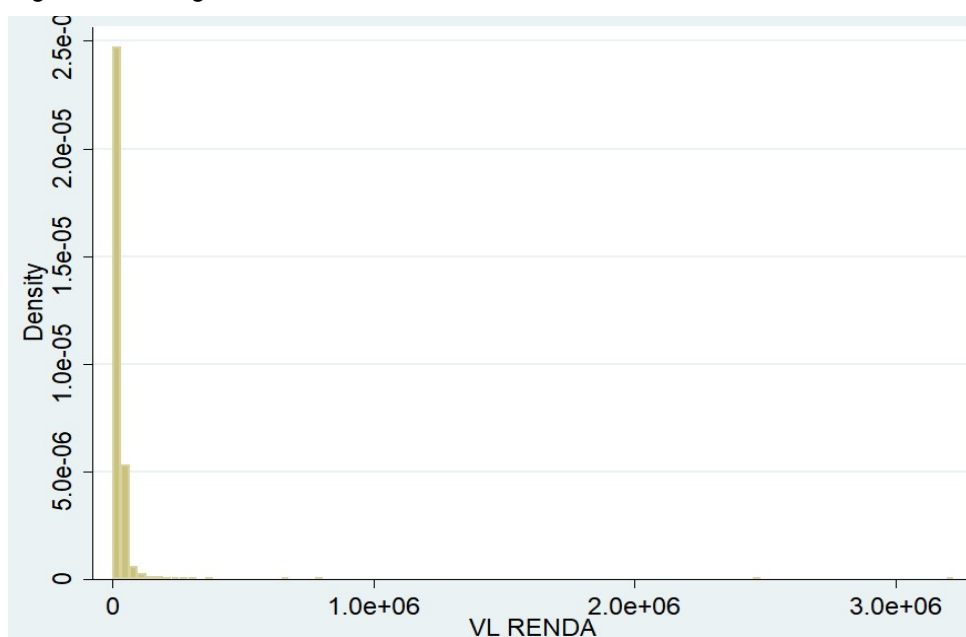
## 4 APRESENTAÇÃO E ANÁLISE DE RESULTADOS

O presente tópico tem por objetivo apresentar os resultados da modelagem de risco de crédito da cooperativa de crédito estudada, a partir da base de dados disponibilizada, com o uso da regressão logística. Serão apresentados os resultados dos procedimentos realizados até a definição dos dois modelos finais, de acordo com a definição de inadimplência, de atraso acima de 30 e 90 dias. Posteriormente serão realizados testes para verificação de eficiência dos modelos.

### 4.1 MODELOS FINAIS

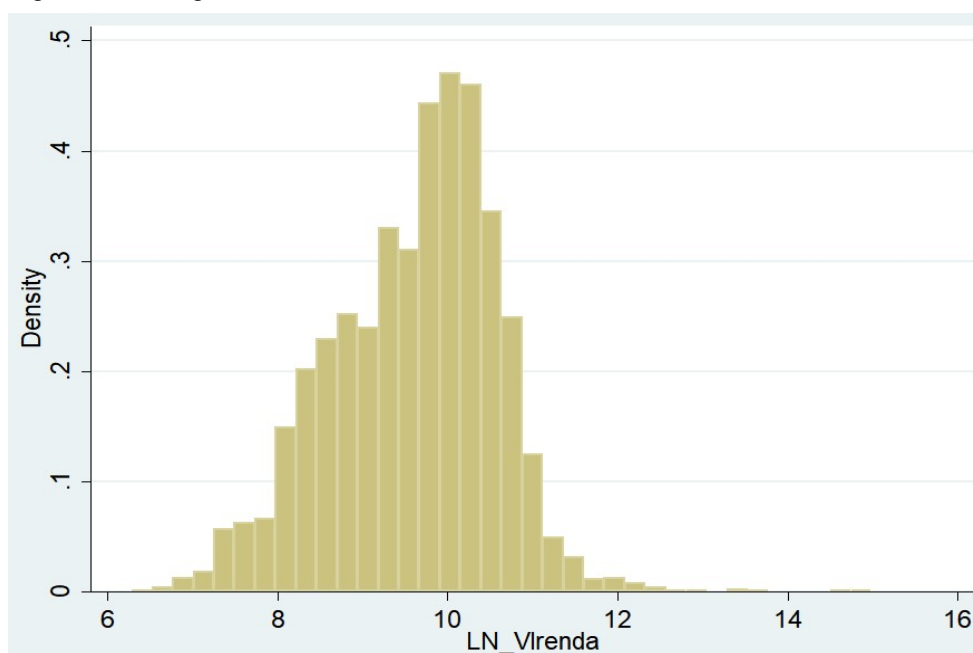
A partir de uma série de procedimentos estatísticos para modelagem, entre eles a visualização gráfica de histogramas, como nas Figuras 4 e 5, foi constatado que, apesar da não necessidade do pressuposto de normalidade das variáveis independentes (HAIR JUNIOR et al., 2009), haveria uma melhora nítida em se utilizar do logaritmo neperiano para as variáveis independentes monetárias (valor solicitado, valor da renda, valor em aplicações, valor em cotas, média de saldo em conta correntes, valor Serasa e saldo devedor), as quais estão particularmente sujeitas a vieses provocados por *outliers*.

Figura 4 - Histograma da variável valor renda



Fonte: resultado da pesquisa, 2018.

Figura 5 - Histograma da variável LN valor renda



Fonte: resultado da pesquisa, 2018.

Conforme descrito anteriormente, algumas variáveis explicativas foram codificadas em *dummies* por se tratarem de variáveis qualitativas. A variável Conta, assumiu os valores 1 para as contas conjuntas e 0 para as contas individuais; Categoria do Empréstimo foi criada duas variáveis, devido a existir 3 categorias, onde se considerou o Crédito sem Garantias como o padrão (0), Financiamento de Bens e Crédito com Garantias Futuras como 1; Sexo assumiu o valor 1 para o sexo masculino e 0 para o sexo feminino; Bens em Garantia e Valor Monetário em Garantia assumiu valores 1, quando as solicitações apresentavam as garantias e 0, quando não as apresentava.

#### 4.1.1 Modelo com definição de inadimplência com atraso maior que 30 dias

O modelo final, para definição de inadimplência de atraso maior que 30 dias, foi estimado através do *software* STATA, conforme citado anteriormente, onde a variável respostas foi classificada como 1, para aquelas propostas aprovadas e com atraso a partir de 31 dias, e 0 para as propostas aprovadas e sem atraso ou com atraso limite de até 30 dias. O modelo possuía um total de 4345 observações, sendo 284 inadimplentes e 4061 adimplentes.

As 20 variáveis explicativas foram incluídas no procedimento *forward stepwise* na regressão logística, e resultaram em 12 variáveis dentro do nível de significância estatística, menor ou igual a 0,05. As variáveis selecionadas para o modelo final constam na Tabela 4. Além disso, pode ser observado os respectivos comportamentos, coeficientes e demais estatísticas de cada variável estimada para o modelo final.

Tabela 4 - Variáveis e coeficientes do modelo final 1

inad30	Coef.	Std. Err.	z	P> z	[95% Conf.	Interval
Indconj	-0.767	0.193	-3.980	0.000	-1.145	-0.390
garfut	-0.983	0.254	-3.870	0.000	-1.481	-0.486
ln_saldev	0.832	0.016	5.130	0.000	0.051	0.115
qtdparc	0.012	0.003	4.320	0.000	0.006	0.017
ln_depcc6m	-0.189	0.022	-8.540	0.000	-0.232	-0.145
anosassoc	-0.056	0.016	-3.410	0.001	-0.088	-0.024
qtdchq12m	0.032	0.010	3.120	0.002	0.121	0.528
idade	0.020	0.006	3.450	0.001	0.009	0.032
limcredito	0.000	0.000	-2.580	0.010	0.000	0.000
ln_vlsolic	0.219	0.066	3.310	0.001	0.089	0.348
ln_vlserasa	0.100	0.021	4.780	0.000	0.059	0.141
ln_deppraz	-0.111	0.032	-3.480	0.000	-0.173	-0.048
_cons	-4.877	0.655	-7.440	0.000	-6.161	-3.593

Fonte: resultados da pesquisa, 2018.

Cita-se novamente que algumas variáveis, usualmente, utilizadas em estimações de *credit score* acabaram não sendo incluídas no trabalho pelos diversos fatores apresentados anteriormente, no tópico 3.3.3.

As 12 variáveis explicativas selecionadas para o modelo final foram:

1. CONTA (Indconj): Variável qualitativa (*dummy*), onde 1 representou ser conta conjunta, ou seja, possuir mais de um titular, e 0 representou conta individual. Demonstrou que contas com mais de um titular, apresentam menor risco de inadimplência;
2. CRÉDITO DE ANTECIPAÇÃO DE GARANTIAS FUTURAS (*garfut*): Variável qualitativa (*dummy*), onde 1 representou se o crédito solicitado era da categoria de créditos solicitados com garantias futuras (antecipação de convênios e recebíveis debitados na cooperativa; antecipação de imposto

de renda a receber na cooperativa; e crédito consignado), e 0 se é da categoria de crédito sem garantias reais (aquisição de cotas em outras cooperativas de ramo diferente; crédito destinado a congressos e treinamento profissional; ação sustentável; crédito pessoal; móveis; reforma residencial e comercial; viagem; IPTU; IPVA; e *hot money*). Apresentou comportamento redutor para a variável resposta, ou seja, se a solicitação de crédito é da categoria crédito de antecipação de garantias futuras, irá apresentar um menor risco;

3. SALDO DEVEDOR (In\_saldev): Variável quantitativa, que representou o total de empréstimos devidos na cooperativa. Quanto maior esse saldo apontou ser, maior o risco de inadimplência;
4. QUANTIDADE DE PARCELAS (qtdparc): Variável quantitativa, que representou a quantidade de parcelas mensais a serem pagas do crédito solicitado. Demonstrou que quanto maior o número de parcelas, maior o risco de inadimplência do solicitante;
5. MÉDIA DE SALDO EM CONTA CORRENTE DOS ÚLTIMOS 6 MESES (In\_depcc6m): Variável quantitativa, que representou o valor médio de saldo em conta corrente nos últimos 6 meses. Apresentou que quanto maior o saldo médio, menor o risco de inadimplência;
6. TEMPO DE RELACIONAMENTO (anosassc): Variável quantitativa, que representou a quantidade de anos que a pessoa é cooperada na cooperativa, chamado tempo de relacionamento nas instituições financeiras. Apontou que quanto maior o tempo de relacionamento com a cooperativa, menor o risco de inadimplência;
7. CHEQUES DEVOLVIDOS NOS ÚLTIMOS 12 MESES (qtdchq12m): Variável quantitativa, que representou a quantidade de cheques devolvidas por insuficiência de saldo nos últimos 12 meses. Demonstrou que quanto maior o número de cheques devolvidos pelo motivo de insuficiência de saldo, maior o risco de inadimplência;

8. IDADE (idade): Variável quantitativa, que representa a idade do cooperado solicitante de crédito. Demonstrou que quanto mais velho, maior o risco de inadimplência;
9. VALOR LIMITE EM CHEQUE ESPECIAL (limcredito): Variável quantitativa, que representa o valor de cheque especial que o cooperado possui na cooperativa. Quanto maior o limite de cheque especial, menor o risco de inadimplência;
10. VALOR SOLICITADO (ln\_vlsolic): Variável quantitativa, que representa o valor do crédito solicitado. Apontou que quanto maior esse valor, maior o risco de inadimplência;
11. CONSULTA SERASA (ln\_vlserasa): Variável quantitativa, que representa o valor de débitos apontado pelo Serasa. Demonstrou que quanto maior o valor apontado, maior o risco de inadimplência;
12. VALOR EM APLICAÇÕES (ln\_deppraz): Variável quantitativa, que representa o valor em aplicações na cooperativa que o cooperado possuía. Quanto maior o saldo em aplicações, menor o risco de inadimplência.

Nota-se que nenhuma das variáveis demonstrou ter algum comportamento incoerente com a prática na análise de crédito, os comportamentos indicaram ser convergentes com o que um analista de crédito julgaria.

Na Tabela 5 a matriz de correlação atesta que não há nenhuma correlação acentuada entre às variáveis explicativas, maiores que 0,9 (HAIR JUNIOR et al., 2009), assim sendo, não havendo nenhum caso de alta colinearidade entre as variáveis independentes, um dos pressupostos necessários para a regressão logística (HAIR JUNIOR et al., 2009).



Tabela 5 - Matriz correlação do modelo 1

	<b>inad30</b>	<b>ldconj</b>	<b>garfut</b>	<b>ln_saldev</b>	<b>qtdparc</b>	<b>ln_depcc6m</b>
<b>inad30</b>	1.000					
<b>ldconj</b>	-0.104	1.000				
<b>garfut</b>	-0.071	-0.005	1.000			
<b>ln_saldev</b>	-0.006	0.138	0.101	1.000		
<b>qtdparc</b>	0.149	-0.013	-0.031	0.014	1.000	
<b>ln_depcc6m</b>	-0.201	0.179	0.006	0.213	-0.051	1.000
<b>anosassoc</b>	-0.136	0.240	0.154	0.358	-0.077	0.347
<b>qtdchq12m</b>	0.087	-0.063	-0.021	0.053	0.002	0.008
<b>idade</b>	-0.006	0.158	0.041	0.159	-0.106	0.122
<b>limcredito</b>	-0.080	0.231	0.116	0.316	-0.025	0.262
<b>ln_vlsolic</b>	0.123	0.049	0.077	0.036	0.624	-0.040
<b>ln_vlserasa</b>	0.133	-0.054	-0.015	0.025	0.013	-0.069
<b>ln_deppraz</b>	-0.089	0.100	-0.081	-0.024	-0.011	0.204
	<b>qtdchq12m</b>	<b>idade</b>	<b>limcredito</b>	<b>ln_vlsolic</b>	<b>ln_vlserasa</b>	<b>ln_deppraz</b>
<b>qtdchq12m</b>	1.000					
<b>idade</b>	-0.010	1.000				
<b>limcredito</b>	-0.002	0.382	1.000			
<b>ln_vlsolic</b>	0.020	0.031	0.196	1.000		
<b>ln_vlserasa</b>	0.206	-0.014	-0.030	0.040	1.000	
<b>ln_deppraz</b>	-0.048	0.039	-0.020	-0.103	-0.060	1.000

Fonte: resultados da pesquisa, 2018.

Outro teste para corroboração da ausência de multicolinearidade é o VIF (*Variance Inflation Factor*), e conforme pode-se constatar no Tabela 6, não apresenta um valor preocupante, quanto a multicolinearidade, que seria próximo ou acima de 10 (HAIR JUNIOR et al., 2009).

Tabela 6 - VIF do modelo 1

<b>Variable</b>	<b>VIF</b>	<b>1/VIF</b>
<b>ln_vlsolic</b>	1.85	0.541
<b>anosassoc</b>	1.83	0.547
<b>qtdparc</b>	1.75	0.571
<b>limcredito</b>	1.58	0.634
<b>idade</b>	1.4	0.716
<b>ln_depcc6m</b>	1.24	0.808
<b>ln_saldev</b>	1.21	0.827
<b>ldconj</b>	1.11	0.902
<b>ln_deppraz</b>	1.08	0.923
<b>ln_vlserasa</b>	1.06	0.945
<b>garfut</b>	1.06	0.948
<b>qtdchq12m</b>	1.05	0.949
<b>Mean VIF</b>	1.35	

Fonte: resultados da pesquisa, 2018.

#### 4.1.2 Medidas de desempenho do modelo um

Para a verificação do poder explicativo do modelo final estimado, foi realizado o teste *Hosmer-Lemeshow*. O teste se consiste em dividir a amostra em dez partes, baseando-se nas probabilidades resultantes do modelo estimado, e a partir de um teste qui-quadrado verificar as diferenças significativas entre as frequências estimadas e esperadas observadas em cada um dos dez grupos. Na Tabela 7 é possível observar os resultados do teste.

Tabela 7 - Teste *Hosmer-Lemeshow* do modelo 1

Group	Prob	Obs_1	Exp_1	Obs_0	Exp_0	Total
1	0.0066	0	1.8	435	433.2	435
2	0.0108	2	3.7	432	430.3	434
3	0.0167	3	5.9	432	429.1	435
4	0.0231	9	8.6	425	425.4	434
5	0.0321	8	11.9	427	423.1	435
6	0.0443	16	16.3	418	417.7	434
7	0.0632	27	23.1	408	411.9	435
8	0.0976	34	34.3	400	399.7	434
9	0.169	71	55.6	364	379.4	435
10	0.9214	114	122.9	320	311.1	434

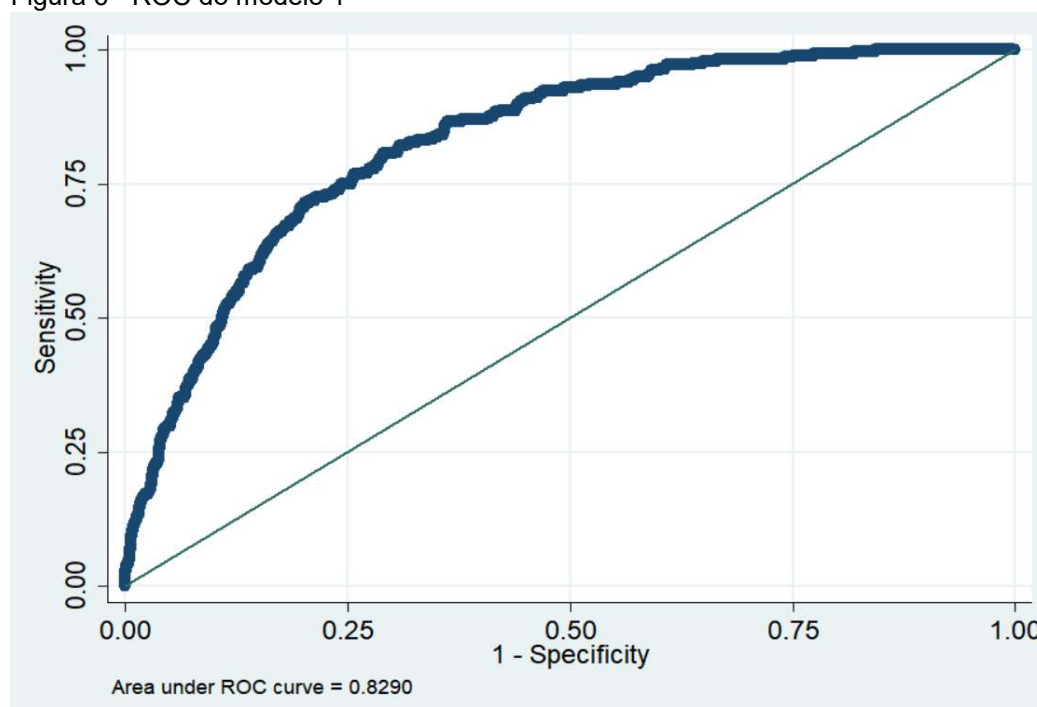
Fonte: resultados da pesquisa, 2018.

O teste apresentou o resultado de 11,92, com oito graus de liberdade e significância de 0,155, indicando que não há diferenças significantes entre os valores reais e observados, assim, se aponta para um ajuste aceitável.

O teste curva ROC (*Receiver Operating Characteristic*), é uma ferramenta usualmente aplicada na verificação da eficiência dos modelos de *credit score*. A curva ROC é um gráfico elaborado a partir da variação da sensibilidade em função de (1 - especificidade). A sensibilidade é o número de acertos, a um determinado *cutoff*, sob os casos evento, ou seja, é o percentual de acerto dentre as observações que de fato são eventos. Já a especificidade tem o mesmo significado, porém dentre as observações não eventos. Através da curva ROC é possível verificar a medida de desempenho AUC (*Area Under the Curve*), que na verdade representa a área total de abrangência abaixo da curva. Quanto mais próximo de 1 o AUC, melhor a eficiência do modelo.

Como é possível verificar a curva de ROC e o resultado da AUC, do modelo final, na Figura 6, o valor da AUC é de 0,829, o que nos remete a um excelente valor para o modelo. Valores acima de 0,9 são classificados como valores acima do comum; maior ou igual a 0,8 até 0,9 são valores excelentes; maior ou igual a 0,7 até 0,8 são valores aceitáveis (SUMIHARA FILHO; SLEEGER, 2010).

Figura 6 - ROC do modelo 1



Fonte: resultado da pesquisa, 2018.

Por fim a matriz de classificação, onde é possível verificar os erros de tipo um e tipo dois, cujo são os percentuais de classificação errônea para os casos de evento classificado como não-evento e os casos de não-evento classificados como evento.

Para elaboração de tal matriz é necessário antes definir o valor de *cutoff*, esse valor definirá a classificação de cada uma das observações em evento ou não evento a partir do resultado gerado no modelo *credit score* estimado. A definição do valor de corte acaba por ser um ponto subjetivo da política de cada empresa na implementação do *credit score*. Ao se aceitar em maior ou menor grau um tipo de erro 1 ou 2, estará influenciando em maior ou menor grau na especificidade ou sensibilidade. Para tal, o trabalho decidiu apresentar valores de corte que, apesar de não apresentarem o valor máximo possível de classificações corretas, esteja maximizando sensibilidade, ou seja, que maximize as classificações de inadimplentes como inadimplentes, ainda que haja prejuízo nas classificações de adimplentes como adimplentes. Porém, sempre foi

observado o objetivo geral da pesquisa que pretende criar uma ferramenta que amplie a automação do processo de análise de crédito a pessoas físicas, ou seja, que por mais que se queira maximizar a sensibilidade, o *cutoff* definido, também deverá trazer uma especificidade com valor consideravelmente acima dos 20%, cujo percentual de aprovações automatizadas atualmente, que se visa aumentar.

A partir dessa explanação, a Tabela 8 apresenta alguns pontos de corte que poderiam ser utilizados na implementação desse modelo de *credit score*, a ser definido pelos gestores da cooperativa.

Tabela 8 - Matriz de classificação do modelo 1

<b>Cutoff - 0.1</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	181	662	843	21.5%
<b>Adimplente</b>	103	3399	3502	97.1%
<b>Total</b>	284	4061	4345	82.4%
<b>Sensitividade</b>	63.7%	<b>Especificidade</b>	83.7%	-
<b>Cutoff - 0.05</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	236	1348	1584	14.90%
<b>Adimplente</b>	48	2713	2761	98.26%
<b>Total</b>	284	4061	4345	67.9%
<b>Sensitividade</b>	83.1%	<b>Especificidade</b>	66.8%	-
<b>Cutoff - 0.03</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	262	1993	2255	11.62%
<b>Adimplente</b>	22	2068	2090	98.95%
<b>Total</b>	284	4061	4345	53.6%
<b>Sensitividade</b>	92.3%	<b>Especificidade</b>	50.9%	-

Fonte: resultados da pesquisa, 2018.

Conforme comentado anteriormente, a definição do *cutoff* dependeria da política a ser decidida pelos gestores da cooperativa. Entretanto, tomando como base os pontos de corte exibidos acima, podemos concluir que as variações da especificidade e sensibilidade acabam tendo comportamento inversamente

proporcional, ou seja, conforme se altera o ponto de corte para cima, e a especificidade aumenta, a sensibilidade acaba por diminuir. Para um *cutoff* de 0,03, possui-se uma especificidade de 50,9% e uma sensibilidade de 92,3%; para 0,05, seria 66,8% e 83,1%, respectivamente; e para 0,1, 83,7% e 63,7%.

Considerando que, se utilizado o modelo de *credit score* na cooperativa para análise e concessão de crédito, as propostas classificadas como inadimplentes iriam ser aprovadas automaticamente e as rejeitadas iriam para a análise no setor de crédito, teríamos os percentuais para um *cutoff* de 0,1, de 82,4% propostas aprovadas automaticamente; para 0,05, de 67,9%; e para 0,03 de 53,6% das propostas, conforme na Tabela 9.

Tabela 9 - Percentual de aprovações automáticas

<b>Cutoff</b>	<b>Aprovadas</b>	<b>%</b>	<b>Rejeitadas</b>	<b>%</b>	<b>Total</b>
<b>0.1</b>	3580	82.4%	765	17.6%	
<b>0.05</b>	2949	67.9%	1396	32.1%	4345
<b>0.03</b>	2330	53.6%	2015	46.4%	

Fonte: resultados da pesquisa, 2018.

Lembra-se, que apesar do *cutoff* de maior valor apresentar um resultado que iria automatizar a análise e concessão de crédito em um grau elevado, iria trazer consigo um menor valor de sensibilidade, que acabaria aumentando a aprovação de propostas inadimplentes.

#### **4.1.3 Modelo com definição de inadimplência com atraso maior que 90 dias**

No modelo final, para definição de inadimplência de atraso maior que 90 dias, onde a variável respostas foi classificada como 1, para aquelas propostas aprovadas e com atraso a partir de 91 dias, e 0 para as propostas aprovadas sem atraso ou com atraso limite de até 90 dias. O modelo possuía um total de 4345 observações, sendo 98 inadimplentes e 4247 adimplentes.

As 20 variáveis explicativas foram incluídas no procedimento *forward stepwise* na regressão logística, e resultaram em 6 variáveis dentro do nível de significância estatística, menor ou igual a 0,05. As variáveis selecionadas para o modelo final constam na Tabela 10 abaixo. Além disso, pode ser observado os respectivos

comportamentos, coeficientes e demais estatísticas de cada variável estimada para o modelo final.

Tabela 10 - Variáveis e coeficientes do modelo final 2

inad90	Coef.	Std. Err.	z	P> z	[95% Conf.	Interval
<b>ln_deppraz</b>	-0.172	0.066	-2.610	0.009	-0.301	-0.043
<b>garfut</b>	-0.922	0.432	-2.130	0.033	-1.770	-0.074
<b>ln_saldev</b>	0.098	0.025	3.950	0.000	0.050	0.147
<b>qtdparc</b>	0.015	0.003	5.310	0.000	0.010	0.021
<b>ln_depcc6m</b>	-0.200	0.035	-5.750	0.000	-0.269	-0.132
<b>anosassoc</b>	-0.107	0.025	-4.300	0.000	-0.156	-0.058
<b>_cons</b>	-2.939	0.240	-12.250	0.000	-3.409	2.469

Fonte: resultados da pesquisa, 2018.

Cita-se novamente que algumas variáveis, usualmente, utilizadas em estimações de credit score acabaram não sendo incluídas no trabalho pelos diversos fatores apresentados anteriormente, no tópico 3.3.3.

As 6 variáveis explicativas selecionadas para o modelo final foram:

1. VALOR EM APLICAÇÕES (ln\_deppraz): Variável quantitativa, que representOU o valor em aplicações na cooperativa que o cooperado possuía. Quanto maior o saldo em aplicações, menor o risco de inadimplência;
2. CRÉDITO DE ANTECIPAÇÃO DE GARANTIAS FUTURAS (garfut): Variável qualitativa (*dummy*), onde 1 representOU se o crédito solicitado era da categoria de créditos solicitados com garantias futuras (antecipação de convênios e recebíveis debitados na cooperativa; antecipação de imposto de renda a receber na cooperativa; e crédito consignado), e 0 se é da categoria de crédito sem garantias reais (aquisição de cotas em outras cooperativas de ramo diferente; crédito destinado a congressos e treinamento profissional; ação sustentável; crédito pessoal; móveis; reforma residencial e comercial; viagem; IPTU; IPVA; e *hot money*). Apresentou comportamento redutor para a variável resposta, ou seja, se a solicitação de crédito é da categoria crédito de antecipação de garantias futuras, irá apresentar um menor risco do que o crédito sem garantias;

3. SALDO DEVEDOR (ln\_saldev): Variável quantitativa, que representou o total de empréstimos devidos na cooperativa. Quanto maior esse saldo apontou ser, maior o risco de inadimplência;
4. QUANTIDADE DE PARCELAS (qtdparc): Variável quantitativa, que representa a quantidade de parcelas mensais a serem pagas do crédito solicitado. Demonstrou que quanto maior o número de parcelas, maior o risco de inadimplência do solicitante;
5. MÉDIA DE SALDO EM CONTA CORRENTE DOS ÚLTIMOS 6 MESES (ln\_depcc6m): Variável quantitativa, que representa o valor médio de saldo em conta corrente nos últimos 6 meses. Apresentou que quanto maior o saldo, menor o risco de inadimplência;
6. TEMPO DE RELACIONAMENTO (anosassc): Variável quantitativa, que representa a quantidade de anos que é cooperado na cooperativa, chamado tempo de relacionamento nas instituições financeiras. Apontou que quanto maior o tempo de relacionamento com a cooperativa, menor o risco de inadimplência.

Nota-se que nenhuma das variáveis demonstrou ter algum comportamento incoerente com a prática na análise de crédito, os comportamentos indicaram ser convergentes com o que um analista de crédito julgaria.

Na Tabela 11 a matriz de correlação atesta que não há nenhuma correlação acentuada, valores acima de 0,9 (HAIR JUNIOR et al., 2009). Assim sendo, não havendo nenhum caso de alta colinearidade, um dos pressupostos necessários para a regressão logística (HAIR JUNIOR et al., 2009).

Tabela 11 - Matriz correlação do modelo 2

	inad90	In_deppraz z	garfut	In_saldev v	qtdparc c	In_depcc6m	anosassoc
inad90	1.000						
In_deppraz	-0.058	1.000					
garfut	-0.043	-0.081	1.000				
In_saldev	0.003	-0.024	0.101	1.000			
qtdparc	0.096	-0.011	-0.031	0.014	1.000		
In_depcc6m	-0.127	0.204	0.006	0.213	-0.051	1.000	
anosassoc	-0.100	0.039	0.154	0.358	-0.077	0.347	1.000

Fonte: resultados da pesquisa, 2018.

Outro teste para corroboração da ausência de multicolinearidade é o VIF (*Variance Inflation Factor*), e conforme pode-se constatar na Tabela 12, não apresenta um valor preocupante, quanto a multicolinearidade, este que seria próximo ou acima de 10 (HAIR JUNIOR et al., 2009).

Tabela 12 - VIF modelo 2

Variable	VIF	1/VIF
anosassoc	1.29	0.776
In_depcc6m	1.2	0.830
In_saldev	1.17	0.855
In_deppraz	1.06	0.948
garfut	1.04	0.965
qtdparc	1.01	0.990
Mean VIF	1.13	

Fonte: resultados da pesquisa, 2018.

#### 4.1.4 Medidas de desempenho do modelo dois

Para a verificação do poder explicativo do modelo final estimado, foi realizado o teste *Hosmer-Lemeshow*. Na Tabela 13 é possível observar os resultados do teste.



Tabela 13 - Teste Hosmer-Lemeshow do modelo 2

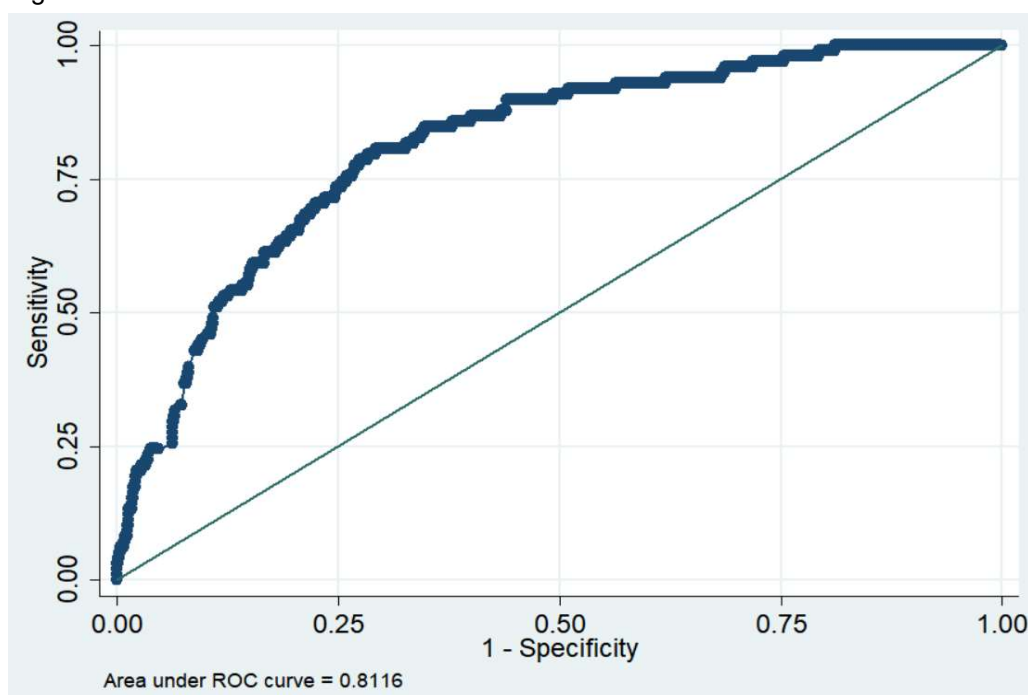
Group	Prob	Obs_1	Exp_1	Obs_0	Exp_0	Total
1	0.0018	0	0.4	435	434.6	435
2	0.0036	1	1.1	433	432.9	434
3	0.0055	3	2	432	433.0	435
4	0.0079	3	2.9	431	431.1	434
5	0.0108	3	4.0	432	431.0	435
6	0.0149	4	5.5	430	428.5	434
7	0.0215	6	7.8	429	427.2	435
8	0.0343	16	11.7	418	422.3	434
9	0.0593	19	19.5	416	415.5	435
10	0.4557	43	42.9	391	391.1	434

Fonte: resultados da pesquisa, 2018.

O teste apresentou o resultado de 3,75, com oito graus de liberdade e significância de 0,879, indicando que não há diferenças significantes entre os valores reais e observados, assim, se aponta para um ajuste aceitável.

O teste curva ROC (*Receiver Operating Characteristic*), é uma ferramenta usualmente aplicada na verificação da eficiência dos modelos de *credit score*. A curva ROC é um gráfico elaborado a partir da variação da sensibilidade em função de (1 - especificidade). A sensibilidade é o número de acertos, a um determinado *cutoff*, sob os casos evento, ou seja, é o percentual de acerto dentre as observações que de fato são eventos. Já a especificidade tem o mesmo significado, porém dentre as observações não eventos. Através da curva ROC é possível verificar a medida de desempenho AUC (*Area Under the Curve*), que na verdade representa a área total de abrangência abaixo da curva. Quanto mais próximo de 1 o AUC, melhor a eficiência do modelo. No Figura 7, é possível verificar a curva de ROC e o resultado da AUC, do modelo final.

Figura 7 - ROC do modelo 2



Fonte: resultado da pesquisa, 2018.

Como verificado, o valor da AUC é de 0,812, o que nos remete a um excelente valor para o modelo estimado, conforme citado anteriormente no tópico 4.1.2 a respeito das faixas de valores da AUC e sua classificação.

Por fim a matriz de classificação, onde é possível verificar os erros de tipo um e tipo dois, cujo são os percentuais de classificação errônea para os casos de evento classificado como não-evento e os casos de não-evento classificados como evento.

Na Tabela 14, é apresentado alguns pontos de corte que poderiam ser utilizados na implementação desse modelo. Porém, sempre foi observado o objetivo geral da pesquisa, que pretende criar uma ferramenta que amplie a automação do processo de análise de crédito a pessoas físicas, ou seja, que por mais que se queira minimizar o erro de tipo um, o *cutoff* apresentado sempre deverá trazer um valor de especificidade consideravelmente acima dos 20% de automação do processo de análise de crédito da cooperativa.

Tabela 14 - Matriz de classificação do modelo 2

<b>Cutoff - 0.05</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	51	498	549	9.3%
<b>Adimplente</b>	47	3749	3796	98.8%
<b>Total</b>	98	4247	4345	87.5%
<b>Sensitividade</b>	52.0%	<b>Especificidade</b>	88.3%	-
<b>Cutoff - 0.025</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	72	1070	1142	6.30%
<b>Adimplente</b>	26	3177	3203	99.19%
<b>Total</b>	98	4247	4345	74.8%
<b>Sensitividade</b>	73.5%	<b>Especificidade</b>	74.8%	-
<b>Cutoff - 0.01</b>				
	<b>Inadimplente</b>	<b>Adimplente</b>	<b>Total</b>	<b>% Acerto</b>
<b>Inadimplente</b>	90	2184	2274	3.96%
<b>Adimplente</b>	8	2063	2071	99.61%
<b>Total</b>	98	4247	4345	49.6%
<b>Sensitividade</b>	91.8%	<b>Especificidade</b>	48.6%	-

Fonte: resultados da pesquisa, 2018.

Conforme explanado anteriormente, a definição do *cutoff*, para implementação, dependeria da política a ser decidida pelos gestores da cooperativa. Entretanto, tomando como base os exemplos citados acima, podemos concluir que as variações da especificidade e sensibilidade acabam tendo comportamento inversamente proporcional, ou seja, conforme se altera o ponto de corte para cima, e a especificidade aumenta, a sensibilidade acaba por diminuir. Para um *cutoff* de 0,01, possui-se uma especificidade de 48,6% e uma sensibilidade de 91,8%; para 0,025, seria 74,8% e 73,5%, respectivamente; e para 0,05, 88,3% e 52%.

Considerando que, se utilizado o modelo de *credit score* na cooperativa para análise e concessão de crédito, as propostas classificadas como adimplentes iriam ser aprovadas automaticamente e as rejeitadas iriam para a análise no setor de crédito, teríamos os percentuais para um *cutoff* de 0,05, de 87,5% propostas aprovadas automaticamente; para 0,025, de 74,8%; e para 0,01 de 49,6% das propostas, conforme na Tabela 15.

Tabela 15 - Percentual de Aprovações Automáticas 2

<b>Cutoff</b>	<b>Aprovadas</b>	<b>%</b>	<b>Rejeitadas</b>	<b>%</b>	<b>Total</b>
<b>0.05</b>	3800	87.5%	545	12.5%	4345
<b>0.025</b>	3249	74.8%	1096	25.2%	
<b>0.01</b>	2153	49.6%	2192	50.4%	

Fonte: resultados da pesquisa, 2018.

Lembra-se, que apesar do *cutoff* de maior valor apresentar um resultado que iria automatizar a análise e concessão de crédito em um grau elevado, iria trazer consigo um menor valor de sensibilidade, que acabaria aumentando a aprovação de propostas inadimplentes.

## 5 CONCLUSÃO

O objetivo do estudo foi estimar um modelo de *credit score*, para a análise e concessão automática de crédito a solicitantes, pessoas físicas, cooperados na cooperativa de crédito X. Através da base de dados disponibilizada foi estudada junto aos funcionários e a própria data base, quais informações teriam fidedignidade para serem utilizadas no desenvolvimento do modelo. Concluiu-se que a base de dados teria um total de 20 variáveis fidedignas, algumas convergentes com a literatura e outras convergentes com a política da empresa, que foram utilizadas para estimação do *credit score*, distribuídas em variáveis de cadastro – oriundas da base cadastral dos cooperados na cooperativa -, variáveis de utilização – oriundas do histórico de uso dos produtos e serviços financeiros utilizados na cooperativa – e variável de restrição – oriunda do sistema Serasa.

Definida a base de dados, optou-se por estimar dois modelos de *credit score*. Um modelo austero, onde se considerou inadimplentes aqueles que incorrem em um atraso de pagamento maior que 30 dias, e outro mais brando, onde se considerou inadimplentes aqueles que incorreram em um atraso de pagamento maior que 90 dias. Além disso, constatou-se na base de dados 62 linhas de créditos diferentes que foram classificadas em 5 categorias, de acordo com a natureza de cada linha. Dessas categorias uma apresentou um número ínfimo de observações para a modelagem e outra apresentava um número irrisório de inadimplência, bem menor ainda que a inadimplência geral da cooperativa, que pode ser constatada, com base nos dados disponibilizados, como uma carteira com baixíssimo nível de inadimplência, e foram retiradas da estimação. Logo os modelos se restringiram a 3 categorias, ou 46 linhas de crédito.

Sobre os modelos, foi estimado um primeiro modelo mais rigoroso, considerando que a definição de inadimplentes eram os empréstimos que demonstraram atraso acima de 30 dias de pagamento. Este modelo estimado selecionou 12 variáveis, como significativas, dentro das 20 iniciais. Todos os comportamentos, dos coeficientes, convergiram com o que se espera de uma avaliação de um analista de crédito, não havendo nenhuma contrariedade empírica com o modelo estimado. Os testes apontaram não haver problemas de multicolinearidade, demonstrou um ajuste aceitável e uma curva ROC de nível excelente, no valor de 0,829.

O segundo modelo, era mais brando, e considerou a inadimplência como atrasos maiores que 90 dias de pagamento. Este modelo selecionou 6 variáveis, como significantes. Novamente, todas as variáveis selecionadas convergiram com comportamentos esperados empiricamente. O modelo se adequou nos testes e apresentou excelente valor de curva ROC, de 0,812.

Diferente de trabalhos anteriores, relacionados com o tema, este decidiu por não apresentar o maior valor de precisão preditiva. A questão de implementação dos modelos passa por fatores gerenciais e de políticas das empresas.

Neste trabalho, foram apresentadas algumas matrizes de classificação, com diferentes pontos de cortes baseando-se na minimização das classificações errôneas de inadimplentes, a maximização de sensibilidade notada seria o fator que tem maior apelo empírico. Muito embora a literatura demonstre importar-se muito com o percentual de precisão das classificações gerais, tanto que diversos trabalhos, em especial os nacionais, acabam por realizarem comparativos de percentuais entre os diversos métodos estatísticos, o fato de o trabalho possuir esse apelo faz com que tal peso não seja considerado essencial. Valeu-se mais em demonstrar que os modelos de *credit score* gerado ampliariam de forma significativa a automação do processo, levando-se em conta uma das principais atenções das instituições financeiras, que é preocupação com a exposição ao risco de crédito. Ainda, dentre essas demonstrações, de três exemplos, é possível notar a subjetividade, mesmo com a atenção dada, para definição do ponto de corte.

Sobre as limitações da pesquisa constituiu-se, à falta de fidedignidade de informações cadastrais, a inexistência de informações no banco de dados da cooperativa e restrições de acesso a algumas informações. Também, o fato de não ter tido acesso a informações importantes convergentes com a literatura e com a prática do mercado, a exemplo as informações do SCR (Sistema de Informação de Crédito) do Banco Central do Brasil - onde seria possível incluir diversas variáveis (Empréstimos com outras IF's; Limites de crédito disponíveis no Sistema Financeiro Nacional; Prejuízos com IF's; etc.) -, o histórico de comportamento de crédito com a cooperativa - onde seria possível ver atrasos e variações de pagamentos de empréstimos tomados no passado na cooperativa -, limites e utilização dos cartões de créditos, estado civil, quantidade de filhos, débitos automáticos na cooperativa, entre outros. Outra limitação foi o problema temporal para a disponibilização de uma base

de dados de validação do modelo, onde seria possível validar os modelos estimados em uma outra base de dados futura (2016 e 2017), não utilizados na estimação.

## REFERÊNCIAS

ALTMAN, E. I.; CAQUETTE, J. B.; NARAYANAN, P. **Managing credit risk - the next great financial challenge**. New York: Hoh Wiley & Son Inc., 1998.

AMARAL, Q. R.; OREIRO, L. J.; PAULA, F. L. Por que as taxas de juros são tão elevadas no Brasil? Uma avaliação empírica. **Revista de Economia Política**, v. 32, n. 4, p. 557-579, out./dez., 2012.

ANDRADE, A. P. S. **Avaliação da recente expansão do crédito no brasil: boom ou crescimento sustentável?**. Dissertação (Mestrado em Administração), São Paulo, Universidade de São Paulo, 2012.

ANIBAL, C. A.; KOYAMA, S. M. Cooperativas de Crédito: taxa de juros praticadas e fatores de viabilidade. **Trabalhos para Discussão – Banco Central do Brasil**, n.257, nov/2011.

ASSAF, A.; BRITO, G. A. S.; CORRAR, L. J. Sistema de classificação de risco de crédito: Uma aplicação a companhias abertas no Brasil. **Revista Contabilidade & Finanças - USP**, v. 20, n. 51, p. 28-43, set./dez., 2009.

BECKMAN, T. N.; **CREDITS AND COLLECTIONS**. New York: Mc Graw-Hill, 1962.

BANCO CENTRAL DO BRASIL. **Cooperativas e crédito e seus impactos sociais**. Disponível em: <[http://www.uniprimedourados.com.br/upload/file/pagina/anexo\\_8\\_file1\\_cooperativas-de-crdito-histria-da-evoluo-normativa-no-brasil.pdf](http://www.uniprimedourados.com.br/upload/file/pagina/anexo_8_file1_cooperativas-de-crdito-histria-da-evoluo-normativa-no-brasil.pdf)>. Acesso em: 10 fev. 2017.

BANCO CENTRAL DO BRASIL. **Resolução 1.559/88**. Brasília: 22/12/1988. Disponível em: <[https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/42089/Res\\_1559\\_v1\\_O.pdf](https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/42089/Res_1559_v1_O.pdf) em 16/02/2017>. Acesso em: 23 fev. 2017.

BANCO CENTRAL DO BRASIL. **Resolução 2.682/99**. Brasília: 21/12/1999. Disponível em: <[http://www.bcb.gov.br/pre/normativos/res/1999/pdf/res\\_2682\\_v2\\_L.pdf](http://www.bcb.gov.br/pre/normativos/res/1999/pdf/res_2682_v2_L.pdf) em 20/02/2017>. Acesso em: 17 fev. 2017.

BANCO CENTRAL DO BRASIL. **Sistema de informação de crédito - SCR**. Disponível em: <[http://www.bcb.gov.br/fis/crc/port/SCR\\_Manual\\_cidadao.pdf](http://www.bcb.gov.br/fis/crc/port/SCR_Manual_cidadao.pdf)>. Acesso em: 15 mai. 2017.

BANCO CENTRAL DO BRASIL. **Relatório de Inclusão Financeira 2015**. Disponível em: <<https://www.bcb.gov.br/Nor/relinconfin/RIF2015.pdf>>. Acesso em: 28 mai. 2017.



Basel Committee on Banking Supervision. **International Convergence of Capital Measurement and Capital Standards: a revised framework comprehensive version** Disponível em: < <https://www.bis.org/publ/bcbs128.pdf>>.

BIALOSKORSKI NETO, S. **Economia e gestão de organizações cooperativas**. 2. ed. São Paulo: Atlas, 2014.

BLATT, A. **Avaliação de risco e decisão de crédito: um enfoque prático**. São Paulo: Nobel, 1999.

CHAIA, A. J. **Modelos de gestão do risco de crédito e sua aplicabilidade ao mercado brasileiro**. Dissertação (Mestrado em Administração), São Paulo, Universidade de São Paulo, 2003.

CONANT. **Principles of money and banking**. Credits y cobros. N.Y.: Minerva, 1967.

CORRAR, L. J et al. **Análise multivariada: para os cursos de administração, ciências contábeis e economia**. São Paulo: Atlas, 2014. 541 p.

COSTA, F. N. Bancos e crédito no Brasil: 1945-2007, História e Economia. **Revista Interdisciplinar**, v. 4, n. 2, 2008. São Paulo, BBS/Terra Comunicação Editorial, 2009.

FÁVERO, L. P. **Análise de dados: modelos de regressão com excel, stata e spss**. 1. ed. Rio de Janeiro: Elsevier. 2015.

FISHER, R. A. The use multiple measurements in taxonomic problems. *Annals of Eugenics...* v. 7. p. 179-188. 1936.

FLACH, P. A. ROC analysis. In: **Encyclopedia of machine learning**. New York: Springer, 2010. p. 869-875.

FONSECA, J. J. S. **Metodologia da pesquisa científica**. Fortaleza: UEC Apostila, 2002.

GARTER, I. R.; MOREIRA, T. B. S.; GALVES, M. Análise do risco setorial como instrumento de controle gerencial em instituições financeiras. **RAM - Revista de Administração Mackenzie**, v. 10, n. 5, p. 107-129, 2008.

GITMAN, L. J. **Princípios de administração financeira**. 7. ed. São Paulo: Hibra, 1997.

GOLDSMITH, R. W. Financial structure and development. **The Economic Journal**, v. 80, n. 318, p. 365-367, 1969.

GONÇALVES, F. O.; JACQUES, E. R. Cooperativas de crédito no Brasil: evolução e impacto sobre a renda dos municípios brasileiros. **Revista Economia e Sociedade**, v. 25, n. 2, p. 489-509, 2016.

GUIMARÃES, A.; MOREIRA, T. B. S. Previsão de insolvência: um modelo baseado em índices contáveis com utilização da análise discriminante. **Revista de Economia Contemporânea**, v. 12, n. 1, p. 151-178, 2008.

GUIMARÃES, I. A.; CHAVES NETO, A. Reconhecimento de padrões: metodologias estatísticas em crédito ao consumidor. **RAE-Eletrônica**, Rio de Janeiro, v. 1, n. 2, p. 1-14, 2002.

HAIR JUNIOR, J. F. et al. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookmann, 2009.

HOSMER, D. W.; LEMESHOW, S. **Applied logistic regression**. New York: J. Wiley, 2004. 383 p.

INTERNATIONAL CO-OPERATIVE ALLIANCE. ICA. Disponível em: <<https://ica.coop/>>. Acesso em: 12 fev. 2017.

KEYNES, J. M. **The general theory of employment, interest, and Money**. London: Palgrave Macmillan, 1936.

KING, R.; LEVINE, R. Finance and growth: Schumpeter might be right. **The Quartely Journal of Economics**, 1993.

LEWIS, E. M. **An introduction to credit scoring**. San Rafael: Fair Isaac and Co., Inc, 1992.

LOPES, I. C. R. Aspectos sobre a legalização das sociedades cooperativas. In: GEDIEL, J. A. P. (Org.). **Os caminhos do cooperativismo**. Curitiba: UFPR, 2001.

LOUZADA-NETO, F.; FERREIRA-SILVA, P. H.; DOMOZ, C. A. R. On impact o disporportional samples in credit scoring models: An application to a Brazilian bank data. **Expert Systems with Application**, v. 39, p. 807-8078, 2012.

MCKINNON, R. **Money and capital in economic devolpment**. v. 2, 1973.

ORGANIZAÇÃO DAS COOPERATIVAS BRASILEIRAS. OCB. Disponível em: <[http://www.ocb.org.br/site/cooperativismo/evolucao\\_no\\_brasil.asp](http://www.ocb.org.br/site/cooperativismo/evolucao_no_brasil.asp)>. Acesso em: 17 fev. 2017.

ORGANIZAÇÃO DAS COOPERATIVAS BRASILEIRAS – ALAGOAS. OCB-AL. Disponível em: <<http://ocb-al.com.br/cooperativismo>>. Acesso em: 22 fev. 2018.

PETIT, L.; DEVEYRAC R. **El credito y la organizacion bancaria**. México: Editorial América, p. 39, 1945.

ROSA, P. T. M. **Modelos de credit scoring: Regressão logística, chaid e real**. Dissertação (Mestrado em Estatística), São Paulo, Universidade de São Paulo, IME/USP, 2000.

SANTOS, J. O.; FAMA, R. Avaliação de aplicabilidade de um modelo de credit scoring com variáveis sistêmicas e não-sistêmicas em carteiras de crédito bancário rotativo de pessoas físicas. **Revista Contabilidade & Finanças - USP**, n. 44, p. 105-117, 2007.

SANTOS, J. O. **Análise de crédito** – empresas e pessoas físicas. 2. ed. São Paulo: Atlas, 2003.

SAUNDERS, A. **Administração de instituições financeiras**. 2. ed., Americana, São Paulo: Atlas, 2000.

SCHERR, F. C. **Modern working capital management**. New Jersey: Prentice-Hall, 1989.

SCHRICKEL, W. K. **Análise de crédito: concessão e gerência de empréstimos**. São Paulo: Atlas, 1995.

SECURATO, J. R. **Crédito: análise e avaliação do risco**. São Paulo: Saint Paul, 2002.

SHARMA S. **Applied multivariate techniques**. Hoboken: John Wiley & Sons; 1996.

SICSU, A. L.; **Credit Scoring**. São Paulo: Blucher, 1999.

SILVA, F. G. A. **Risco de crédito bancário e informação assimétrica: teoria e evidência**. Dissertação (Mestrado em Administração), Porto Alegre, Universidade Federal do Rio Grande do Sul, 2004.

SILVA, P. J. **Gestão e análise de risco de crédito**. 5. ed. São Paulo: Atlas. 2006.

SILVA, R. A. **Performance financeira da carteira na avaliação de modelos de análise e concessão de crédito: uma abordagem baseada em aprendizagem estatística**. Tese (Doutorado em Administração), Ribeirão Preto, Universidade São Paulo, 2014.

SOARES, C. de A. F. **Da possibilidade de regulação das agências de rating no Brasil**. Monografia. Rio de Janeiro. Instituto de Economia da Universidade Federal do Rio de Janeiro, 2005.

STEINER, M. T. A. et al. Extração de regras de classificação a partir de redes neurais para auxílio à tomada de decisão na concessão de crédito bancário. **Pesquisa Operacional**, v. 27, n. 3, p. 407-426, 2007.

SUMIHARA FILHO, H.; SLEEGERS, L. C. Valores de referência para os principais indicadores de acurácia dos modelos de escoragem. **Serasa Experian - Tecnologia de Crédito**, São Paulo, n. 73, p. 31-45, ago. 2010.

TAVARES, R.; **Operações de Crédito: produtos e serviços bancários**. Curitiba: Intersaberes, 2014.

VICENTE, E. F. R. **A estimativa do risco na constituição da PDD**. Dissertação (Mestrado em Administração), São Paulo; Universidade de São Paulo, 2001.

WESTON, J. F.; BRIGHAM E. F. **Managerial finance**. New York: Holt, 1972.