

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Departamento de Estatística

Sérgio Antônio Valadares Garcia Pereira

**APLICAÇÃO DO MODELO DE REGRESSÃO LOGÍSTICA NA ANÁLISE DE
INADIMPLÊNCIA EM CONTRATOS DE FINANCIAMENTO HABITACIONAL COM
RECURSOS DO SBPE**

Belo Horizonte
2022

Sérgio Antônio Valadares Garcia Pereira

**APLICAÇÃO DO MODELO DE REGRESSÃO LOGÍSTICA NA ANÁLISE DE
INADIMPLÊNCIA EM CONTRATOS DE FINANCIAMENTO HABITACIONAL COM
RECURSOS DO SBPE**

Monografia de especialização apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, para a obtenção do título de Especialista em Estatística.

Orientadora: Profa. Lourdes Coral Contreras Montenegro

Belo Horizonte
2022

2022, Sérgio Antônio Valadares Garcia Pereira.

Todos os direitos reservados.

Pereira, Sérgio Antônio Valadares Garcia

P436a Aplicação do modelo de regressão logística na análise de inadimplência em contratos de financiamento habitacional com recursos do SBPE [manuscrito] / Sérgio Antônio Valadares Garcia Pereira —2022.

41.f. il.

Orientador: Lourdes Coral Contreras Montenegro.

Monografia (especialização) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística.

Referências: 41

1. Estatística. 2. Inadimplência (Finanças). 3. Inflação. 4. Poupança. 5. Regressão logística. 6. Fundo de garantia do tempo de serviço. I. Contreras Montenegro, Lourdes Coral. II. Universidade Federal de Minas Gerais I. Instituto de Ciências Exatas, Departamento de Estatística. III. Título.

CDU 519.2 (043)

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende Costa CRB 6/1510

Universidade Federal de Minas Gerais – ICEX



Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística
Programa de Pós-Graduação / Especialização
Av. Pres. Antônio Carlos, 6627 - Pampulha
31270-901 – Belo Horizonte – MG

E-mail: pgest@ufmg.br
Tel: 3409-5923 – FAX: 3409-5924

ATA DO 257^a. TRABALHO DE FIM DE CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA DE SÉRGIO ANTÔNIO VALADARES GARCIA PEREIRA.

Aos treze dias do mês de dezembro de 2022, às 14:00 horas, com utilização de recursos de videoconferência a distância, reuniram-se os professores abaixo relacionados, formando a Comissão Examinadora homologada pela Comissão do Curso de Especialização em Estatística, para julgar a apresentação do trabalho de fim de curso do aluno **Sérgio Antônio Valadares Garcia Pereira**, intitulado: “Aplicação do modelo de Regressão Logística na análise de inadimplência em contratos de financiamento habitacional com recursos do SBPE”, como requisito para obtenção do Grau de Especialista em Estatística. Abrindo a sessão, a Presidente da Comissão, Professora Lourdes Coral Contreras Montenegro – Orientadora, após dar conhecimento aos presentes do teor das normas regulamentares, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do candidato. Após a defesa, os membros da banca examinadora reuniram-se sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foi atribuída a seguinte indicação: o candidato foi considerado Aprovado condicional às modificações sugeridas pela banca examinadora no prazo de 30 dias a partir da data de hoje por unanimidade. O resultado final foi comunicado publicamente ao candidato pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente Ata, que será assinada por todos os membros participantes da banca examinadora. Belo Horizonte, 13 de dezembro de 2022.

Prof.^a Lourdes Coral ~~Contreras~~ Montenegro (Orientadora)
Departamento de Estatística / ICEx / UFMG

Jussiane Nader
Gonçalves

Assinado de forma digital por
Jussiane Nader Gonçalves
Dados: 2022.12.14 10:02:45 -03'00'

Prof.^a Jussiane Nader Gonçalves
Departamento de Estatística / ICEx / UFMG

Prof. Guilherme Lopes de Oliveira
DECOM / CEFET-MG

Resumo

Pela lei nº 4.380 de 21 de agosto de 1964 foi criado o Sistema Financeiro de Habitação, SFH. Este novo modelo de financiamento tinha à época como fontes de recursos a poupança e o FGTS, Fundo de Garantia do Tempo de Serviço. O objetivo da criação do SFH é o de reduzir o déficit habitacional no país e de melhorar as formas de financiamento de imóveis. O SFH, como qualquer programa de governo, é impactado diretamente por mudanças nas políticas econômicas, sociais e pelas oscilações de mercado, que podem impactar diretamente na renda das pessoas que possuem financiamento imobiliário, ou mutuários, propiciando um cenário de inadimplência. Este trabalho tem como objetivo apresentar um modelo de regressão logística que estima a probabilidade de um contrato ser inadimplente no financiamento imobiliário com base nas variáveis relacionadas ao perfil do mutuário, no imóvel e no crédito.

Palavras-chave: SFH. SBPE. FGTS. Inadimplência. Inflação. Poupança. Regressão logística.

Abstract

Under the Law No. 4,380 of August 21, 1964 created the Housing Financial System, SFH. At the time, this new financing model had savings and the FGTS, the Severance Indemnity Fund, as sources of funds. The objective of creating the SFH is to reduce the housing deficit in the country and to improve the ways of financing real estate. The SFH, like any government program, is directly impacted by changes in economic and social policies and by market fluctuations, which can directly impact the income of people who have real estate financing, or borrowers, providing a scenario of default. This paper aims to present a logistic regression model that estimates the probability of a contract defaulting on real estate financing based on variables related to the borrower's profile, property and credit.

Keywords: SFH. SBPE. FGTS. Default. Inflation. Savings. Logistic Regression.

LISTA DE FIGURAS

Figura 1- Visualização da regressão logística	24
Figura 2- Gráfico da área sob a curva ROC	26
Figura 3- Curva ROC do modelo selecionado.....	38

LISTA DE TABELAS

Tabela 1 - Estatística descritiva referente às variáveis Renda Familiar e Idade dos mutuários adimplentes e inadimplentes	29
Tabela 2 - Estatística descritiva referente às variáveis Quota e Comprometimento de Renda dos mutuários adimplentes e inadimplentes	30
Tabela 3 - Análise comparativa das variáveis do perfil do mutuário entre grupos.....	31
Tabela 4 - Análise comparativa da variável relacionada à dívida entre grupos.....	32
Tabela 5 - Análise comparativa das variáveis relacionadas ao Imóvel entre grupos	33
Tabela 6 - Resultados do ajuste do modelo de regressão logística considerando as dez variáveis	34
Tabela 7 - Resultados do ajuste do modelo de regressão logística considerando as oito variáveis	35
Tabela 8 - Apresentação dos valores VIF apurados com as oito variáveis selecionadas	36
Tabela 9 - Apresentação das razões de chance (OR) e intervalos de 95% de confiança do modelo final.	37

SUMÁRIO

1. Introdução	10
Objetivo do Trabalho	11
Estrutura do Trabalho.....	11
2. Revisão da Literatura	12
2.1 Sistema Financeiro da Habitação (SFH)	12
2.2 Banco Nacional da Habitação (BNH)	15
2.3 Sistema Brasileiro de Poupança e Empréstimo (SBPE).....	16
2.4 Fundo de Garantia do Tempo de Serviço (FGTS).....	17
2.5 Fundo de Compensação de Variações Salariais (FCVS).....	17
2.6 Inadimplência	18
3. Materiais e Métodos	20
3.1 Base de Dados.....	20
Quadro 1 – Variáveis relacionadas ao mutuário	21
Quadro 2 – Variáveis relacionadas à dívida	21
Quadro 3 – Variáveis relacionadas ao comprometimento de renda	22
3.2 Métodos.....	22
3.2.1 Teste Qui-Quadrado	22
3.2.2 Amostragem Aleatória Simples (AAS) para Proporção.....	22
3.2.3 Regressão logística	23
3.2.4 Curva ROC	25
3.2.5 <i>Softwares</i>	26
4. Análise dos Resultados	27
4.1 Apuração do tamanho da amostra	27
4.2 Análise descritiva dos dados	28
4.3 Análise Comparativa dos Dados	30
4.4 Ajuste do modelo de regressão logística.....	34
5 Considerações Finais	39
6 Referências	41

1. Introdução

O Sistema Financeiro da Habitação, SFH, criado durante o governo do presidente Castelo Branco em 1964 e implementado em 1966, é o plano de habitação mais duradouro e efetivo a propiciar aquisição de moradia à parcela mais carente da população brasileira.

A fim de garantir o sucesso das políticas habitacionais, os governos devem estar atentos à arrecadação e ao controle da inadimplência:

[...] o desempenho do SFH dependeria fundamentalmente de dois fatores básicos: a capacidade de arrecadação do FGTS e do SBPE e o grau de inadimplência dos mutuários. Em outras palavras, essa dependência significava que, apesar da sofisticação do seu desenho, o SFH, como de resto qualquer sistema de financiamento de longo prazo, era essencialmente vulnerável a flutuações econômicas que afetassem estas variáveis (SANTOS, Cláudio Hamilton M., 1999).

Desde sua implantação foram vários os desafios enfrentados pelos governos que se sucederam para manter esta política habitacional, tendo, como último grande desafio, a recente pandemia do Corona vírus¹, em 2020.

Em momentos de crise a tendência é de que a renda das pessoas seja afetada, propiciando um cenário mais favorável à inadimplência. A fim de se antecipar a momentos críticos, que possam corroborar com o aumento de inadimplência, é fundamental que existam mecanismos que auxiliem os agentes financeiros a tentar minimizar a possibilidade de inadimplência nas suas carteiras.

A utilização da regressão logística para predição da inadimplência ou insolvência se iniciou com Ohlson (1980), que desenvolveu o primeiro modelo de regressão logística para prever insolvência de empresas. A partir desse trabalho, a regressão logística passou a ser um método muito utilizado por pesquisadores; como, por exemplo, Camargos *et al.* (2010), avaliaram a situação de crédito, adimplente ou inadimplente, em uma amostra de 9.232 empresas que obtiveram crédito em uma instituição financeira pública de Minas Gerais, no período de junho de 1997 a dezembro de 2005.

¹ <https://www.paho.org/pt/covid19/historico-da-pandemia-covid-19>

Objetivo do Trabalho

Este trabalho tem como objetivo construir um modelo que possa determinar a probabilidade de um contrato vir a ficar inadimplente em um contrato de financiamento imobiliário. Este resultado irá subsidiar, principalmente, setores de avaliação de risco.

Estrutura do Trabalho

O trabalho está dividido em quatro capítulos. O Capítulo 1 faz uma introdução ao projeto a ser descrito nesta monografia. O Capítulo 2 abrange todos os conceitos necessários para um melhor entendimento do projeto. O Capítulo 3 apresenta os materiais e os métodos utilizados neste trabalho. O Capítulo 4 apresenta os resultados obtidos. No Capítulo 5 são apresentadas as conclusões do trabalho e as considerações finais.

2. Revisão da Literatura

Nesta seção será mostrado uma breve história do SFH, contextualizando o leitor em relação ao tema abordado. Seguindo-se a este breve histórico, estarão relacionados os conceitos de alguns dos principais elementos que o compõem.

Com o objetivo principal de reduzir o déficit habitacional no Brasil, foi criado em 21 de agosto de 1964, através da lei nº 4380, o SFH e o BNH, Banco Nacional da Habitação, em meio a uma reformulação geral do Sistema Financeiro Nacional. O BNH era o responsável por gerir o SFH e os recursos do FGTS, Fundo de Garantia do Tempo de Serviço. No SFH, as principais fontes de recurso são o FGTS e os recursos oriundos da poupança, com regras para sua utilização normatizadas pelo SBPE, Sistema Brasileiro de Poupança e Empréstimo.

2.1 Sistema Financeiro da Habitação (SFH)

Uma das principais características do SFH é a possibilidade de se financiar até 80% do imóvel, com até 35 anos para conclusão do pagamento; a depender da idade do mutuário, a possibilidade de utilização do FGTS na operação e a taxa de juros efetiva limitada a 11,3865% a.a., além de ser uma transação de empréstimo isenta da cobrança de IOF (Imposto sobre Operações Financeiras).

Desde a sua criação, como é possível de se verificar no site² da Associação Brasileira das Entidades de Crédito Imobiliário e Poupança (Abecip), o SFH enfrentou diversos altos e baixos.

Da criação até meados da década de 1970, marcado como a fase áurea do financiamento habitacional, com grande captação de recursos em poupança e no FGTS, resultado do aumento no nível de emprego e da massa salarial no país. Neste período foi criado o FCVS, Fundo de Compensação de Variações Salariais, que garantia o saldo residual do contrato após o término do prazo contratual. O objetivo do FCVS era corrigir o

² Abecip. Disponível em: <<https://www.abecip.org.br/credito-imobiliario/historia#:~:text=Sistema%20de%20Financiamento%20Imobili%C3%A1rio,geral%20do%20Sistema%20Financeiro%20Nacional>>.

descompasso entre os períodos de aplicação da correção monetária aos saldos e às prestações.

No final dos anos 1970 e parte da década de 80 acontece o estouro da inflação, em decorrência da segunda crise do petróleo e da dívida externa. As ações judiciais se multiplicaram, visando a adequação das prestações dos empréstimos habitacionais à evolução dos salários, uma vez que a inflação causou um aumento abrupto das parcelas, não tendo os salários acompanhado o mesmo ritmo. As diversas tentativas do Governo de reduzir a inflação, além de não surtirem efeito, pioravam a situação. Com o início do Plano Cruzado em março de 1986, as prestações foram convertidas de cruzeiro para cruzado, levando-se em consideração a média dos últimos doze meses. A partir desse plano, vários outros foram editados: Plano Cruzado II, Plano Bresser, Plano Verão, Plano Collor I e Plano Collor II, sempre congelando preços e salários.

Ainda em 1986 foi extinto o BNH e suas atribuições passaram a ser desenvolvidas pelo Banco Central do Brasil, pela Caixa Econômica Federal e pelo Ministério da Fazenda. Neste mesmo período vários dos contratos concedidos com prazos de 15 a 20 anos foram finalizados, deixando tanto os bancos como o Governo com um enorme rombo, o do FCVS. Nos contratos sem a cobertura do FCVS, a dívida passou a ser desestimulante para que este continuasse adimplente. Em várias ocasiões a dívida ultrapassava o valor de mercado do imóvel.

Década de 90. O início deste período foi marcado pela retenção de boa parte dos recursos de poupança, impactando seriamente a construção civil e o Crédito Imobiliário. O resultado disso foi o esgotamento dos recursos do SFH e fechamento das carteiras habitacionais dos Agentes Financeiros em boa parte da década de 90, contribuindo para um aumento significativo no déficit habitacional no país.

Somente em 1996 a Caixa Econômica Federal voltou a operar com Crédito Habitacional à pessoa física.

Em face aos altos desequilíbrios no SFH, houve a necessidade do surgimento de sistemas alternativos, como o Sistema Financeiro Imobiliário (SFI), lei 9.514/97, instituindo a alienação fiduciária, em alternativa à garantia hipotecária.

Nos anos 2000, o SFH experimentou um retorno aos bons tempos. Com a redução nas taxas de juros e a baixa nos índices de inflação. As empresas e as famílias passaram a assumir compromissos de mais longo prazo.

O SBPE tornou-se principal fonte de recursos para novos financiamentos da classe média. Os recursos do FGTS e de outros fundos sociais desempenharam importante papel no financiamento habitacional para famílias de baixa renda, com destaque para o programa Minha Casa Minha Vida, criado em 2009.

Dos anos 2010 até os dias atuais, destaca-se neste período a crise de 2015, com aumento expressivo da inflação, aumento no número de desempregados e queda na renda. Esta crise teve como fonte o grande endividamento das famílias com pagamento de dívidas. Para se ter uma ideia, em 2005 menos de 18% da renda das famílias estava comprometida com pagamento de dívidas, em 2013 este percentual já era de 45% e, em 2015, chegou a 61,6%. Em 2013, a situação do mercado imobiliário começou a se agravar, em decorrência do aumento da inflação. Em 2014, a fim de evitar um descontrole sobre a inflação, o Governo aumentou a taxa SELIC, aumentou os juros imobiliários, reduziu o crédito para aquisição de imóveis usados e aumentou o valor de entrada para estes tipos de imóveis. Neste cenário ocorreu uma grande fuga de recursos da poupança, reduzindo ainda mais a oferta de crédito. Essa alteração na política de financiamento de imóveis usados prejudicou em muito a comercialização de imóveis novos, uma vez que grande parte das famílias precisa da venda do seu imóvel para adquirir outro, o que resultou numa redução na atividade da construção civil, conforme observou o macroeconomista Mario Jorge Cardoso de Mendonça, do IPEA, em 2013:

De acordo com os resultados um choque contracionista de política monetária produz o efeito negativo sobre todas as variáveis ligadas ao setor imobiliário. Assim, observa-se a queda acentuada e contínua da demanda por crédito imobiliário, do preço dos imóveis e do produto da construção civil assim como aumento da inadimplência. Isso pode ser um indício de que um aumento na taxa de juros pode provocar uma forte retração no mercado imobiliário no Brasil.

Essa crise durou até 2017, mostrando sinais de recuperação ao longo de 2019. Porém, com a chegada da pandemia do Corona vírus ao Brasil, houve uma grande alteração nos planos de crescimento do mercado, com a paralisação de muitos negócios e de vários setores da economia.

Durante a pandemia, ao contrário de todas as expectativas, o mercado imobiliário se mostrou bastante resiliente. As medidas do Governo neste setor, como, por exemplo, a oferta de carência por até 6 meses, contribuíram para estabilizar as taxas de inadimplência. Com a necessidade de ficar em casa e com o trabalho remoto, contribuiu-se também para a estabilidade em todo o setor, inclusive com aumento na demanda, valorizando os imóveis. Outro fator que contribuiu para o enfrentamento da pandemia pelo setor imobiliário foi uma maior disponibilidade de crédito para financiamento. Em 2020, houve um crescimento de 57,5% nos valores financiados em comparação com o ano anterior, segundo a Associação Brasileira das Entidades de Crédito Imobiliário e Poupança (Abecip) em 27/01/2021³.

2.2 Banco Nacional da Habitação (BNH)

O banco deveria ser o gestor e financiador de uma política destinada a, como diz a lei de 1964, “promover a construção e aquisição da casa própria, especialmente pelas classes de menor renda”, bem como a ampliar as oportunidades de emprego e a dinamizar o setor da construção civil. A pedra angular do BNH era o Sistema Financeiro da Habitação (SFH), que tinha por finalidade principal prover recursos que garantissem a execução do Plano Nacional da Habitação⁴.

Todas as operações de financiamento eram regidas pelo mecanismo da correção monetária e calculadas a partir de Unidades Padrão de Capital (UPC). A estrutura interna do banco era composta por uma presidência e diretores de seis grandes áreas, a saber:

³ Abecip. Disponível em: <<https://www.abecip.org.br/admin/assets/uploads/anexos/data-abecip-2020-12.pdf>>.

⁴ Banco Nacional de Habitação (BNH). In: FGV. Disponível em: <<https://www18.fgv.br/cpdoc/acervo/dicionarios/verbete-tematico/banco-nacional-da-habitacao-bnh>>. Acesso em: 04 dez. 2022.

área de Natureza Social; área de Administração e Controle Operacional; área de Operações Especiais, Desenvolvimento Urbano e Apoio; área de Programas Habitacionais; área de Agentes Financeiros, e área de Planejamento, Coordenação, Pesquisa, Controle Programático e Processamento de Dados.

Após 22 anos de existência, o BNH foi extinto pelo decreto de lei nº 2.291, de 21 de dezembro de 1986, do então presidente da República José Sarney, que também transferiu a função de coordenador do SFH para a Caixa Econômica Federal e a de regulador para o Banco Central.

2.3 Sistema Brasileiro de Poupança e Empréstimo (SBPE)

O SBPE foi criado em 1964 tendo como integrantes as instituições financeiras especializadas na concessão de financiamentos habitacionais. O BNH era o órgão regulador e fiscalizador das instituições que compunham o SBPE⁵.

Trata-se de uma linha de financiamento ofertada para pessoas que pretendem comprar imóveis. Vale lembrar que bens imóveis envolvem uma variedade de tipos, como casas, apartamentos e terrenos (urbanos ou rurais).

O SBPE não se restringe ao financiamento imobiliário voltado para casas, apartamentos e outros imóveis residenciais, ou seja, é possível também investir em aluguel para imóveis comerciais, como escolas, hospitais, empresas e outros tipos de negócios.

O SBPE também permite a aquisição de imóveis novos e usados, bem como a compra de terrenos.

De acordo com dados do portal do FGTS⁶, quase R\$ 770 bilhões era o saldo, em março de 2022, do SBPE⁷, principal fonte de recursos para o financiamento imobiliário no país, que em 2021 respondeu por 44% do total de empréstimos no país.

⁵ SBPE. In: Blog Direcional. Disponível em: <https://direcional.com.br/blog/financas/sbpe/#O_que_e_SBPE>. Acesso em: 04 dez. 2022.

⁶ <https://www.fgts.gov.br/Pages/sou-empregador/o-que.aspx>

⁷ <https://portal.loft.com.br/sbpe/>

2.4 Fundo de Garantia do Tempo de Serviço (FGTS)

O Fundo de Garantia do Tempo de Serviço foi criado a partir da lei nº 5.107, de 13 de setembro de 1966 e vigente a partir de 01 de janeiro de 1967, para proteger o trabalhador demitido sem justa causa. O FGTS é constituído de contas vinculadas, abertas em nome de cada trabalhador, quando o empregador efetua o primeiro depósito. O saldo da conta vinculada é formado pelos depósitos mensais efetivados pelo empregador, acrescidos de atualização monetária e juros⁸.

O FGTS tem sido a maior fonte de recursos para Habitação Popular e Saneamento Básico.

2.5 Fundo de Compensação de Variações Salariais (FCVS)

O FCVS é um fundo público de natureza contábil e financeira, criado no âmbito do Sistema Financeiro da Habitação a partir da resolução de nº 25, de 16 de junho de 1967, do Conselho de Administração do extinto Banco Nacional da Habitação (BNH), e ratificado pela lei nº 9.443, de 14 de março de 1997, com a finalidade principal de garantir o limite de prazo para amortização da dívida dos mutuários decorrentes de financiamentos habitacionais⁹.

O fundo era composto por uma taxa paga mensalmente pelos mutuários que optaram por esta modalidade, o Coeficiente de Equiparação Salarial (CES). Este fundo foi importante para garantir a quitação da dívida pelo mutuário ao término do prazo contratual, no qual o resíduo era quitado pelo FCVS. Este resíduo era resultado do desequilíbrio entre a correção do Saldo Devedor e da prestação, gerando um valor excedente a ser pago pelo mutuário, mesmo com o término do prazo contratual.

A trajetória do FCVS apresenta-se intrinsecamente relacionada às políticas de fomento à habitação popular.

⁸ FGTS. Disponível em: <<https://www.fgts.gov.br/Pages/sou-empregador/o-que.aspx>>. Acesso em: 04 dez 2022.

⁹ <https://fundosdegoverno.caixa.gov.br/sicfg/fundos/FCVS/detalhe/sobre/>

2.6 Inadimplência

A inadimplência é o não pagamento de uma conta ou dívida. Assim, de acordo com Maria Piedade Morais (2001), o consumidor Inadimplente é aquele que está com uma dívida em aberto¹⁰.

Neste estudo foram considerados como contratos inadimplentes aqueles que possuem mais de 59 dias em atraso. Este período foi considerado em alinhamento com o Sistema de risco do Banco Central, SCR, em que, contratos com atraso superior a 59 dias, não devem ter considerada nenhuma receita, estando com sua dívida congelada, sem evolução contábil, conforme indicado pelo Bacen, Banco Central: “deve representar a soma do valor presente de cada uma das parcelas da operação vencíveis no período, desconsiderados receitas e encargos de qualquer natureza relativos a atraso igual ou superior a 60 dias.” (Caixa Econômica Federal, 2004). O texto indica que os valores de dívida a serem considerados para envio ao SCR e ao Sistema de Informações de Crédito, do Bacen, devem ser aqueles com valores evoluídos até o 59º dia em atraso.

O SFH, em meados da década de 1980, passou a viver crises complexas e sucessivas, desencadeadas pela deterioração da economia nacional. Uma parte da população ficou desempregada e passou a ter dificuldades de arcar com o pagamento das prestações dos financiamentos habitacionais.

A sucessão de choques heterodoxos na economia e políticas monetárias restritivas desorganizou e influenciou substancialmente a política habitacional vigente, limitando de forma efetiva o acesso ao crédito imobiliário. Isso prejudicou toda uma cadeia composta por indústria da construção civil, mutuários, agentes financeiros e, conseqüentemente, postulantes a uma linha de crédito habitacional, provocando uma maximização no déficit habitacional brasileiro (Morais, 2001).

Os choques provocados pelos diversos planos econômicos, mesmo com todas as medidas adotadas para socorrer e, teoricamente, garantir um retorno ao sistema, não estancaram o quadro da inadimplência.

¹⁰ Inadimplência. Disponível em: <https://www.scielo.br/j/ram/a/tW8TTtjhmNM8t4d959k6mbJ/?lang=pt>

O fenômeno da inadimplência contaminou todo o sistema, prejudicou o retorno do capital investido e, conseqüentemente, impediu a concessão de novos financiamentos.

Segundo o boletim de 2004 da Caixa Econômica Federal, “em 1985, as prestações dos contratos habitacionais regidos pelo Plano de Equivalência Salarial por Categoria Profissional (PES/CP), por força de dispositivos legais, contaram com reajuste de 112% em detrimento de 246% que foi a real variação do período”.

Em 1986, a economia brasileira viveu sob a égide do Plano Cruzado, que adotou o mecanismo de conversão da moeda. Para o cálculo das prestações habitacionais, considerou-se a média das 12 últimas prestações, desprezando o histórico do contrato. Depois de adotado esse mecanismo, implementou-se o congelamento das prestações por mais 12 meses.

No ano seguinte, 1987, o Brasil viveu o Plano Bresser, que adotou a conversão das prestações habitacionais pela Unidade de Referência de Preços (URP). Um novo congelamento das prestações é trazido em 1990 por força de um novo choque na economia, o Plano Collor.

Essas ocorrências, dentre outras, causaram impacto frontal ao desenvolvimento dos contratos habitacionais e, conseqüentemente, à vida dos seus titulares, principalmente por elas terem gerado saldos devedores consideráveis, desequilibrando financeiramente os contratos e, em várias situações, com o agravante de o valor de mercado do imóvel ser bem inferior ao da sua dívida junto aos agentes financeiros.

Outro grave momento conjuntural foi o período da pandemia em 2020, no qual, o governo agiu antecipadamente, evitando impactos significativos nos números de inadimplência no SFH, garantindo a prorrogação do pagamento das parcelas do financiamento.

Em 2022 a inadimplência está controlada, muito em decorrência da manutenção de medidas adotadas durante o período emergencial da Covid, a exemplo da carência de pagamento de parcelas nos primeiros meses de assinatura, beneficiando especialmente os mutuários da classe média, consumidores principais de produtos que contam com recursos do SBPE.

3. Materiais e Métodos

3.1 Base de Dados

As informações utilizadas neste trabalho foram extraídas de uma base de dados completa de financiamentos com origem de recurso SBPE. Foi feita a seleção dos contratos na situação em que se encontravam em agosto de 2019, período este imediatamente anterior à pandemia de Covid-19.

A base de dados conta com 887.150 contratos em evolução. Deste total de contratos, 38.860 estavam inadimplentes, considerando o critério de atraso superior a 59 dias e, por óbvio, o restante, ou 848.290 sendo considerados como adimplentes. Neste trabalho, foram considerados contratos inadimplentes aqueles com atraso superior a 59 dias, o que corresponde ao período de atraso em relação à prestação mais antiga em aberto.

Em observância à Lei Geral de Proteção de Dados (LGPD), correspondente à lei nº 13.853, de 8 de julho de 2019, foram omitidas informações relacionadas à origem dos dados.

Para a seleção do tamanho de amostra, será utilizada a metodologia de Amostragem Aleatória Simples (AAS) com reposição para proporções, considerando-se neste trabalho, a proporção de inadimplentes e adimplentes.

Neste contexto, foi analisada a variável-resposta como inadimplente e variáveis independentes que abrangem os principais componentes relacionados ao financiamento: ao mutuário, à dívida e ao imóvel, sendo assim dividido:

- Variáveis relacionadas ao mutuário: Renda, Renda Informal, Idade, Sexo, Estado Civil;
- Variáveis relacionadas à dívida: Situação, Quota, Mais de um participante, Comprometimento de Renda;
- Variáveis relacionadas ao imóvel: Região, Categoria Imóvel.

Os quadros (1 ao 3) apresentam as características das variáveis relacionadas ao mutuário, à dívida e ao imóvel, respectivamente.

Quadro 1 – Variáveis relacionadas ao mutuário

Variável	Tipo	Descrição
Renda Familiar	quantitativa contínua	indica o valor da renda familiar em número de salários mínimos (s.m.). Corresponde ao somatório das rendas de todos os participantes do contrato, sendo composta de renda informal, fora da carteira de trabalho, e formal.
Renda Informal	qualitativa nominal	indica se parte da renda do mutuário, ou sua totalidade, é composta de renda informal, sendo seu domínio: SIM, caso possua, ou NÃO.
Idade	quantitativa discreta	corresponde à idade do mutuário em anos, sendo que, caso o contrato possua mais de um participante, o atributo representa a idade do mutuário mais velho.
Sexo	qualitativa nominal	representa o gênero do mutuário principal. Domínio da variável: Masculino, Feminino.
Estado Civil	qualitativa nominal	demonstra o estado civil do mutuário principal. Domínio da variável: Solteiro, Casado, Demais.

Fonte: elaborado pelo autor

Quadro 2 – Variáveis relacionadas à dívida

Variável	Tipo	Descrição
Situação	qualitativa nominal	indica se o contrato está inadimplente ou não, sendo seu domínio: ADIMPLENTE ou INADIMPLENTE.
Quota	quantitativa contínua	representa, em percentual, quanto do valor do imóvel foi financiado, ou seja, quanto maior a quota, maior o valor da dívida em relação ao valor do imóvel.
Mais de um participante	qualitativa nominal	indica se no contrato existe mais de um participante que compõe a renda familiar informada ou não, sendo seu domínio: SIM e NÃO.
Comprometimento de Renda	quantitativa contínua	representa quanto do valor da Renda Familiar está comprometido com o pagamento da dívida. Este valor tem correspondência direta com o valor do financiamento adquirido.

Fonte: elaborado pelo autor

Quadro 3 – Variáveis relacionadas ao comprometimento de renda

Variável	Tipo	Descrição
Região	qualitativa nominal	demonstra a região geográfica do imóvel financiado. Domínio do atributo: SUL, SUDESTE, CENTRO OESTE, NORDESTE ou NORTE.
Categoria	qualitativa nominal	determina se o imóvel financiado é <i>novo</i> ou <i>usado</i> . É considerado Imóvel Novo se ele não foi transacionado (comprado ou vendido) e se está na primeira locação.

Fonte: elaborado pelo autor

3.2 Métodos

Nesta seção será apresentada as metodologias utilizadas neste trabalho, principalmente, na parte da amostragem e do modelo estatístico.

3.2.1 Teste Qui-Quadrado

É um teste de hipóteses que se destina a encontrar um valor da dispersão para duas variáveis categóricas nominais e avaliar a associação existente entre variáveis qualitativas¹¹.

3.2.2 Amostragem Aleatória Simples (AAS) para Proporção

De acordo com Bolfarine e Bussab (2005), em muitas situações, existe interesse em estudar a proporção de elementos em certa população que possuem determinada característica, como, por exemplo, um item pode ser ou não defeituoso, ser ou não eleitor de determinado partido político e assim por diante. Destas situações, a cada elemento da população está associada a variável

$$Y_i = \begin{cases} 1, & \text{se o elemento } i \text{ possui a características} \\ 0, & \text{caso contrário} \end{cases}$$

Assim,

¹¹

Teste

Qui-Quadrado.

Disponível

em:

<http://www.leg.ufpr.br/lib/exe/fetch.php/disciplinas:ce001:teste_do_qui-quadrado.pdf>

$$P = \frac{1}{N} \sum_{i=1}^N Y_i$$

é a proporção de unidades na população que possuem a característica de interesse. A partir de Bolfarine e Bussab (2005), se pode mostrar que o valor do tamanho da amostra (n), em uma AAS para proporção, utiliza-se a seguinte expressão

$$n = \frac{PQ}{D} \quad (1)$$

onde,

P = proporção de sucessos;

$Q = 1 - P$;

$D = \frac{B^2}{z_{\alpha/2}^2}$, onde B é o erro máximo de estimação da amostra, sendo B fixado, com determinado grau de confiança e $z_{\alpha/2}$ é de uma distribuição normal padrão que limita uma área de $\alpha/2$ acima dele.. O valor de B é definido, geralmente, pelo pesquisador.

Após o cálculo do tamanho da amostra foi realizada a análise descritiva dos dados quantitativos e aplicado o teste Qui-quadrado em relação às variáveis qualitativas e a variável dependente.

3.2.3 Regressão logística

Esta metodologia descreve a relação entre diversas variáveis independentes (X_1, \dots, X_{p-1}), que podem ser categóricas ou não, e uma variável dependente (Y), porém seu resultado é binário, assumindo apenas valores 1 ou 0, consistindo na ocorrência ou não do evento que se deseja analisar. Desta forma, o modelo logístico é dado por:

$$P(Y = 1) = \frac{e^\eta}{1+e^\eta} \quad (2)$$

em que $\eta = \beta_0 + \sum_{i=1}^{p-1} \beta_i X_i$, para $\eta \in (-\infty, \infty)$, $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}$ são os parâmetros que serão estimados pelo método de máxima verossimilhança e X_1, X_2, \dots, X_{p-1} correspondem às variáveis independentes que compõe o modelo¹².

¹² Regressão logística. Disponível em: <<https://www.tibco.com/pt-br/reference-center/what-is-logistic-regression#:~:text=A%20regress%C3%A3o%20log%C3%ADstica%20%C3%A9%20usada,valor%20%C3%A9%201%20ou%200.>>>. Acesso em: 04 dez 2022.

Existem três tipos básicos de regressão logística (explicadas através de exemplos):

Regressão logística binária:

Você votou na última eleição?	0 – Não 1 – Sim
-------------------------------	--------------------

Regressão logística multinomial:

Você concorda ou discorda do presidente?	1 – Discordo 2 – Neutro 3 – Concordo
--	--

Regressão logística ordinal:

Qual o grau de instrução?	1 – Básico 2 – Médio 3 – Graduado
---------------------------	---

No presente trabalho será utilizada a regressão logística binária, que a partir de agora, será chamada apenas de regressão logística. Neste trabalho, foi considerada como variável dependente se um contrato está adimplente ou não.

A curva logística representa a relação entre as variáveis, dependente e independentes, desta forma os valores previstos ficam entre 0 e 1, sendo definido pelos coeficientes estimados.

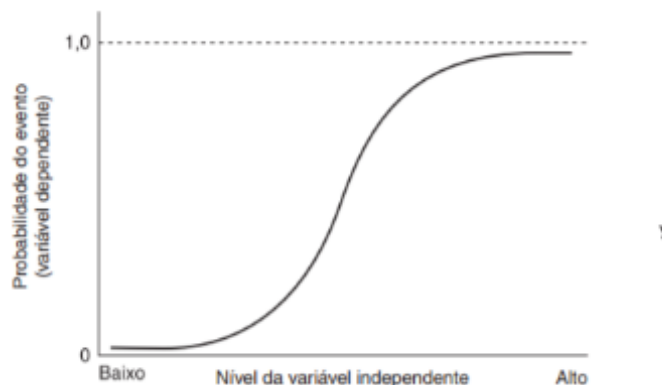


Figura 1-Visualização da regressão logística

Fonte: <https://statplace.com.br/blog/modelos-de-machine-learning-utilizando-o-pacote-caret/>

3.2.4 Curva ROC

A curva ROC, ou *Receiver Operating Characteristic*, tem por objetivo mensurar a capacidade de predição do modelo proposto, através das predições de sensibilidade e especificidade.

Segundo Fawcett (2006), esta técnica serve para visualizar, organizar e classificar o modelo com base na performance preditiva.

A sensibilidade é a probabilidade de o modelo estimar um resultado positivo em comparação ao resultado real positivo (*Taxa de verdadeiro positivo*).

A especificidade é a probabilidade de o modelo estimar um resultado negativo em comparação ao resultado real negativo (*Taxa de falso positivo*).

No gráfico da curva ROC o eixo *y* representa a *Taxa de verdadeiro positivo* e o eixo *x* a *Taxa de falso positivo*, sendo $1 - \text{Especificidade}$. A capacidade de predição do modelo pode ser avaliada em relação à área sob a curva ROC, ou AUC, na sigla em inglês, *Area Under the ROC Curve*.

O valor de AUC varia de 0 a 1. Dessa forma, quanto maior o AUC melhor a capacidade preditiva do modelo.

Hosmer e Lemeshow (2000) sugerem a utilização de AUC acima de 0,7 como aceitável.

A geração da curva ROC foi gerada a partir do software R, utilizando comandos do pacote pROC¹³.

¹³<https://cran.r-project.org/web/packages/pROC/pROC.pdf>

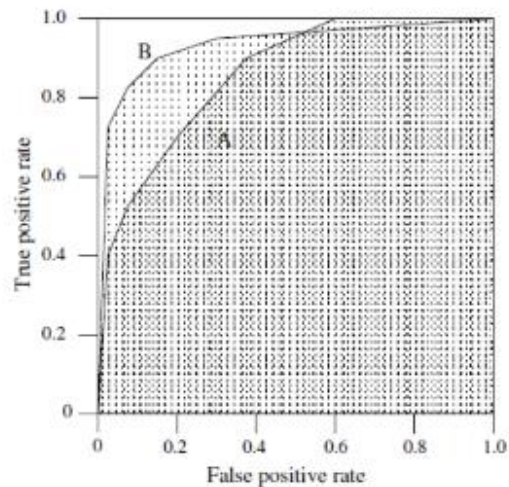


Figura 2- Gráfico da área sob a curva ROC

Fonte: <https://staesthetic.wordpress.com/2014/04/14/gini-roc-auc-and-accuracy/>

Na Figura 2 pode-se verificar que a área sob a curva B é maior que sob a curva A, desta forma o modelo da curva B possui melhor poder preditivo que o modelo representado pela curva A.

3.2.5 Softwares

RStudio

Para a realização dos métodos aqui apresentados foi utilizado o *software* livre RStudio¹⁴, que é uma interface gráfica para desenvolvimento integrado na linguagem R¹⁵, sendo este uma linguagem de programação que tem por finalidade a programação estatística e gráfica, especializada na manipulação, análise e visualização de dados.

¹⁴ RStudio. In: Posit. Disponível em: <<https://posit.co/products/open-source/rstudio/>>. Acesso em: 04 dez. 2022.

¹⁵ Linguagem R. In: Blog Betrybe. Disponível em: <<https://blog.betrybe.com/linguagem-de-programacao/linguagem-r-tudo-sobre/>>. Acesso em: 04 dez. 2022.

Jamovi

O Jamovi¹⁶ é uma ferramenta recente para aplicações estatísticas, desenvolvida para ser uma alternativa barata a outros *softwares* estatísticos como Minitab, SPSS, SAS, entre outros.

4. Análise dos Resultados

Nesta seção, será apresentado o cálculo do tamanho da amostra, os resultados da análise descritiva dos grupos inadimplentes e adimplentes e os resultados do teste de Qui-quadrado. Em seguida se ajustou o modelo de regressão logística das variáveis que foram selecionadas no teste, sendo selecionado aquele modelo com o menor AIC, Critério de Informação de Akaike e, finalmente, gerada a curva ROC para conhecer a consistência da predição dos dados.

4.1 Apuração do tamanho da amostra

Utilizando a Equação (1), foram calculados os termos de cada componente da equação, considerando o nível de confiança de 95%.

$$P = \frac{38.860}{887.150} = 0,0438,$$

$$Q = 1 - 0,0438 = 0,9562,$$

$$D = \frac{(0,003)^2}{(1,96)^2} = \frac{0,000009}{3,8416} = 2,3427 * 10^{-6},$$

Portanto,

$$n = \frac{0,9562 * 0,0438}{2,3427 * 10^{-6}} = 17.877$$

Para este trabalho foi definido o valor do erro amostral (B) de 0,003, com a finalidade de alcançar um número significativo de contratos inadimplentes.

O tamanho da amostra apurada é de 17.877 contratos. A seleção dos contratos foi realizada por sorteio com igual probabilidade de seleção de cada um dos contratos, sem

¹⁶ JAMOVI. In: Prof. Cremona. Disponível em: <<https://www.profcremona.com/academia/posgraduacao/jamovi>>. Acesso em: 04 dez. 2022.

a retirada do mesmo para a seleção do contrato seguinte, isto representa, uma AASc, até completar os totais definidos.

Portanto, a amostra referente aos contratos inadimplentes é de 783 contratos e, para os contratos adimplentes, 17.094 contratos.

Após a seleção foram excluídos alguns dos registros por apresentarem variáveis com valor nulo. Desta forma, após o tratamento dos dados, a amostra final apresentou um total de 17.087 contratos adimplentes e 748 contratos inadimplentes.

Como o número de contratos excluídos da amostra, 0,23% do total, foi mínimo, e dado o tamanho da amostra remanescente, optei pela não geração de uma nova amostra.

4.2 Análise descritiva dos dados

A Tabela 1 apresenta as estatísticas descritivas da Renda Familiar e Idade dos mutuários adimplentes e inadimplente. Para o grupo adimplentes é possível notar que as médias são: (i) Renda Familiar de R\$ 15,00 s.m.; e (ii) 45,40 anos de idade. Os mutuários adimplentes possuem Renda Familiar mínima de R\$ 1,50 s.m. e, máxima, de R\$750,00 s.m., a idade mínima do mutuário é de 19 anos e, máxima, de 83 anos. Avaliando o Coeficiente de Variação (CV) é possível verificar que a Renda Familiar possui maior variabilidade comparando com a variável idade.

Tabela 1 - Estatística descritiva referente às variáveis Renda Familiar e Idade dos mutuários adimplentes e inadimplentes

Adimplentes			Inadimplentes		
Medidas	Renda Familiar	Idade	Medidas	Renda Familiar	Idade
Média	15,00	45,40	Média	16,00	45,20
Mediana	11,80	44,00	Mediana	12,10	44,00
Desvio			Desvio		
Padrão	14,10	10,01	Padrão	21,30	9,52
Mínimo	1,50	19	Mínimo	2,95	24
Máximo	750,00	83	Máximo	528,00	76
CV(%)	94,00	22,05	CV(%)	133,13	21,06

Fonte: elaborado pelo autor

Para o grupo inadimplentes apresentam médias de (i) Renda Familiar de R\$ 16,00 s.m. e (ii) 45,40 anos de idade. Os mutuários inadimplentes possuem Renda Familiar mínima de R\$ 2,95 s.m. e máximo de R\$528,00 s.m., a idade mínima do mutuário é de 24 anos e máximo de 75 anos. Avaliando o CV é possível verificar que, assim como os contratos adimplentes, a Renda Familiar possui variabilidade muito maior em relação à variável idade. Porém, a idade média entre os grupos adimplente e Inadimplente apresentaram resultados próximos. Vale notar ainda que a variável Renda dos mutuários inadimplentes possui variabilidade maior do que em relação aos mutuários adimplentes.

A Tabela 2 apresenta as estatísticas descritivas da Quota e Comprometimento de Renda para os grupos de adimplentes e inadimplentes. É possível notar que as médias são (i) Quota de 72,30% e (ii) 16,10% de Comprometimento de Renda. Os mutuários adimplentes possuem Quota mínima de 0,05% e máxima de 100,00%, o Comprometimento de Renda mínimo é de 0,15% e máximo de 73,00%. Avaliando o Coeficiente de Variação (CV) é possível verificar que a Quota possui menor variabilidade comparando com a variável Comprometimento de Renda.

Tabela 2 - Estatística descritiva referente às variáveis Quota e Comprometimento de Renda dos mutuários adimplentes e inadimplentes

Adimplentes			Inadimplentes		
Medidas	Quota	Comprometimento da Renda	Medidas	Quota	Comprometimento da Renda
Média	72,30	16,10	Média	80,00	21,50
Mediana	77,90	16,40	Mediana	85,00	21,70
Desvio Padrão	17,70	7,24	Desvio Padrão	13,20	9,73
Mínimo	0,05	0,15	Mínimo	17,40	0,46
Máximo	100,00	73,00	Máximo	100,00	97,10
CV(%)	24,48	44,97	CV(%)	21,75	45,26

Fonte: elaborado pelo autor

Para o grupo inadimplentes, observa-se que a Quota média de financiamento do valor do imóvel para o grupo Inadimplentes é de 80% em relação ao valor do imóvel e o valor médio de Comprometimento da Renda dos mutuários é de 21,50%. Além disso, a Quota mínima é de 17,40% e máxima de 100%, o mínimo de Comprometimento de Renda é de 0,46% e, máximo, de 97,10%. Avaliando o CV, é possível verificar que, assim como no estrato de adimplentes, a Quota possui variabilidade menor em relação à variável Comprometimento de Renda. Vale notar ainda que a variável Comprometimento de Renda para o grupo Inadimplentes possui variabilidade maior do que em relação ao estrato de adimplentes

4.3 Análise Comparativa dos Dados

Nas tabelas desta seção são apresentados os resultados do teste Qui-Quadrado das variáveis independentes qualitativas em relação à variável dependente. Foi utilizado um nível de significância de 10% para validar uma associação como significativa com a variável-resposta. Verifica-se que, dentre o total da amostra de 17.835 contratos, 4,20% destes (n=748) foram incluídos no grupo de inadimplentes e os demais, 95,80% (n=17087), no estrato dos adimplentes.

Na Tabela 3 observa-se que, entre os grupos, a maioria dos contratos, 66,31%, são compostos por mutuários do sexo masculino para o grupo Inadimplentes e 63,19% para os adimplentes. Em relação ao Estado Civil do mutuário nota-se que os casados apresentam 45,19% entre os inadimplentes e 50,73% entre os adimplentes. O segundo Estado Civil que se destaca é de solteiro, com maior porcentagem (40,64%) para inadimplentes do que para o grupo adimplentes (35,04%).

Tabela 3 - Análise comparativa das variáveis do perfil do mutuário entre grupos

Características	Total de Mutuários		inadimplentes				Valor-p
			Sim		Não		
	17.835	100%	748	4,20%	17.087	95,80%	
Sexo							
Feminino	6.541	36,70%	252	33,69%	6.289	36,81%	0,083
Masculino	11.294	63,30%	496	66,31%	10.798	63,19%	
Estado Civil							0,005
Casado	8.669	50,50%	338	45,19%	8.669	50,73%	
Demais	2.431	14,20%	106	14,17%	2.431	14,23%	
Solteiro	5.987	35,30%	304	40,64%	5.987	35,04%	
Renda Informal							< 0,001
Não	15.459	86,70%	588	78,61%	14.871	87,03%	
Sim	2.376	13,30%	160	21,39%	2.216	12,97%	

Fonte: elaborado pelo autor

Em relação aos contratos que não possuem renda informal na composição da Renda Familiar, verifica-se que, para o grupo adimplentes, 87,03% dos mutuários não possuem Renda Informal, do que se comparado com o grupo Inadimplentes, que corresponde a 78,61% dos mutuários que não possuem Renda Informal. Diferente do que se apura em

relação aos mutuários com Renda Informal, na qual o número de inadimplentes é maior (21,39%) que o de adimplentes (12,97%).

Em relação a associação das variáveis do perfil do mutuário com inadimplência quando aplicado o teste de significância (valor-p $\leq 0,10$), pode-se observar que as variáveis Sexo, Estado Civil e Renda Informal são significativas a 10% de significância. Portanto, essas variáveis foram selecionadas para fazer parte do modelo de regressão logística. A variável Renda Informal apresentou o menor valor-p, o que pode ser entendido que mutuários que possuem apenas renda formal (carteira assinada), se caracterizam por apresentarem menor possibilidade de inadimplência.

Na Tabela 4 observa-se entre os grupos que os contratos com Mais de um Participante igual a Sim são maioria, com aproximadamente 63,50%. Nos contratos com Mais de um Participante igual a Não, verifica-se uma maior ocorrência de contratos adimplentes do que inadimplentes, 63,95% contra 53,34%, enquanto nos contratos com apenas um participante observa-se um percentual de 46,66% de inadimplentes contra 36,05% dentre os adimplentes.

Tabela 4 - Análise comparativa da variável relacionada à dívida entre grupos

Características	Total de Mutuários		inadimplentes				Valor-p
			Sim		Não		
	17.835	100%	748	4,20%	17.087	95,80%	
Mais de um Participante							
Não	6.509	36,50%	349	46,66%	6.160	36,05%	< 0,001
Sim	11.326	63,50%	399	53,34%	10.927	63,95%	

Fonte: elaborado pelo autor

Em relação a associação da variável Mais de um Participante com variável dependente correspondente aos inadimplentes, quando aplicado o teste de significância (valor-p $\leq 0,10$), pode-se observar que a variável é significativa a 10% de significância. Portanto, essa variável foi selecionada para fazer parte do modelo de regressão logística.

A Tabela 5 apresenta a análise comparativa entre os grupos Inadimplente e adimplente quanto às variáveis Categoria e Região. Ao analisar a variável Categoria observa-se que 71,12% do grupo adimplentes financia imóveis usados e 69,92% para o grupo inadimplentes. Observe-se também que a maioria dos financiamentos se concentra na Região Sudeste, sendo 49,60% para o grupo inadimplentes e, 50,49%, para os adimplentes. A segunda Região que se destaca é o Nordeste que apresenta uma maior porcentagem no grupo dos contratos inadimplentes.

Tabela 5 - Análise comparativa das variáveis relacionadas ao Imóvel entre grupos

Características	Total de Mutuários		inadimplentes				Valor-p
			Sim		Não		
	17.835	100%	748	4,20%	17.087	95,80%	
Categoria							
Novo	5.160	28,93%	225	30,08%	4.935	28,88%	0,479
Usado	12.675	71,07%	523	69,92%	12.152	71,12%	
Região							
Centro-Oeste	1.790	10,04%	66	8,82%	1.724	10,09%	< 0,001
Nordeste	2.814	15,78%	160	21,39%	2.654	15,53%	
Norte	666	3,73%	49	6,55%	617	3,61%	
Sudeste	8.999	50,46%	371	49,60%	8.628	50,49%	
Sul	3.566	19,99%	102	13,64%	3.464	20,27%	

Fonte: elaborado pelo autor

Em relação a associação das variáveis relacionada ao imóvel, verifica-se que apenas a variável Região possui associação significativa a um nível de 10% com inadimplência em relação a variável resposta, enquanto a variável Categoria, imóvel novo ou usado, não é determinante para compor o modelo de inadimplência, uma vez que o estado do imóvel, como novo ou usado. Desta forma, apenas a variável Região será incluída no modelo de regressão logística, tendo como referência a Região Sul uma vez que, proporcionalmente dentro do grupo de regiões, foi o que apresentou, proporcionalmente, a menor inadimplência.

4.4 Ajuste do modelo de regressão logística

Na análise descritiva dos dados e da aplicação do teste Qui-quadrado, foram consideradas dez variáveis. Dentre as dez variáveis analisadas, oito foram selecionadas como sendo candidatas a variáveis explicativas para fazer parte do ajuste do modelo de regressão logística. As variáveis são as seguintes: Renda Familiar, Quota, Sexo, Comprometimento de Renda, Estado Civil, Renda Informal, Mais de um Participante e Região.

Antes do ajuste do modelo com as 8 variáveis selecionadas, foi feito o ajuste do modelo com todas as dez variáveis, gerando um AIC de 3900,1, Tabela 6, superior ao AIC da Tabela 7, de 3896,3, ratificando a exclusão das variáveis Idade e Categoria Imóvel.

A Tabela 6 apresenta o resultado do ajuste do modelo com todas as variáveis, mostrando os coeficientes para cada uma delas.

Tabela 6 - Resultados do ajuste do modelo de regressão logística considerando as dez variáveis

Variável	Coeficiente	Desvio Padrão	Valor Z	Valor-P
(Intercepto)	-7,862633	0,433047	-18,157	< 0,001
Idade	-0,001544	0,005125	-0,301	0,763
Categoria Imóvel Usado	0,023741	0,101679	0,233	0,815
Renda	0,008178	0,001866	4,383	< 0,001
Sexo Masculino	0,390574	0,103133	3,787	< 0,001
Quota	0,029843	0,003668	8,135	< 0,001
Comprometimento Renda	0,090999	0,006679	13,625	< 0,001
Estado Civil Solteiro	-0,351693	0,165808	-2,121	0,033
Estado Civil Demais	-0,473083	0,201421	-2,349	0,018
Renda Informal sim	0,564686	0,116193	4,860	< 0,001
Mais de um Participante não	0,730271	0,160641	4,546	< 0,001
Região Centro Oeste	0,038256	0,191727	0,200	0,841
Região Nordeste	0,519715	0,154715	3,359	< 0,001
Região Norte	0,879632	0,217670	4,041	< 0,001
Região Sudeste	0,107216	0,133988	0,800	0,423
AIC = 3900,1				

Fonte: elaborado pelo autor

Conforme observado na Tabela 6 as variáveis Idade e Categoria Imóvel apresentaram valores-p muito superiores ao nível de significância de 5%, corroborando com a análise descritiva realizada que sugeria a exclusão destas variáveis. Desta forma será considerado o ajuste do modelo sem estas variáveis.

A Tabela 7 apresenta o resultado do ajuste do modelo mostrando os coeficientes para cada uma das variáveis selecionadas a partir da análise descritiva.

Tabela 7 - Resultados do ajuste do modelo de regressão logística considerando as oito variáveis

Variável	Coeficiente	Desvio Padrão	Valor Z	Valor-P
(Intercepto)	-7,922670	0,348122	-22,758	< 0,001
Renda	0,008147	0,001868	4,362	< 0,001
Sexo Masculino	0,391203	0,103135	3,793	< 0,001
Quota	0,029914	0,003662	8,169	< 0,001
Comprometimento Renda	0,090951	0,006673	13,629	< 0,001
Estado Civil Solteiro	-0,341122	0,161670	-2,110	0,035
Estado Civil Demais	-0,477642	0,200835	-2,378	0,017
Renda Informal sim	0,566026	0,116128	4,874	< 0,001
Mais de um Participante não	0,727558	0,160222	4,541	< 0,001
Região Centro Oeste	0,038281	0,191732	0,200	0,842
Região Nordeste	0,517092	0,154507	3,347	< 0,001
Região Norte	0,877195	0,217591	4,031	< 0,001
Região Sudeste	0,108850	0,133855	0,813	0,416
AIC = 3896,3				

Fonte: elaborado pelo autor

As variáveis Região Centro-Oeste e Região Sudeste não diferem da região de referência (Sul) ao nível de significância de 5%.

Para saber se há multicolinearidade entre as variáveis, foi utilizado o índice de VIF (fator de inflação da variância). A Tabela 8, apresenta os valores do VIF relativos as oito variáveis. Verifica-se que não foi obtido nenhum valor VIF acima de 10, sugerindo que todas as variáveis não são correlacionadas.

Tabela 8 - Apresentação dos valores VIF apurados com as oito variáveis selecionadas

Variáveis	VIF
Renda Familiar	1,09
Sexo	1,07
Quota	1,02
Comprometimento de Renda	1,09
Estado Civil	2,94
Renda Informal	1,05
Mais de um Participante	2,91
Região	1,07

Fonte: elaborado pelo autor

Assim, o modelo de regressão logística ajustado é determinado por:

$$\begin{aligned} \text{logit}\{P(Y = 1)\} = & -7,922670 + 0,008147 * \text{Renda} + 0,391203 * \text{Sexo Masculino} + 0,029914 * \text{Quota} \\ & + 0,090951 * \text{Comprometimento de Renda} - 0,341122 * \text{Estado Civil Solteiro} - 0,477642 \\ & * \text{Estado Civil Demais} + 0,566026 * \text{Renda Informal Sim} + 0,727558 \\ & * \text{Mais de um Participante não} + 0,038281 * \text{Região Centro Oeste} + 0,517092 * \text{Região Nordeste} \\ & + 0,877195 * \text{Região Norte} + 0,108850 * \text{Região Sudeste} \end{aligned}$$

Na Tabela 9 são apresentadas as razões de chance (OR) para cada uma das variáveis do modelo final e seus respectivos intervalos de 95% de confiança.

Tabela 9 - Apresentação das razões de chance (OR) e intervalos de 95% de confiança do modelo final.

Variáveis explicativas	OR	2,50%	97,50%	Interpretação(%)
Renda	1,0082	1,0044	1,0121	0,82
Sexo Masculino	1,4788	1,2105	1,8140	47,88
Quota	1,0304	1,0231	1,0379	3,04
Comprometimento Renda	1,0952	1,0811	1,1098	9,52
Estado Civil Solteiro	0,7110	0,5140	0,9692	-28,90
Estado Civil Demais	0,6202	0,4150	0,9123	-37,98
Renda Informal sim	1,7613	1,3976	2,2041	76,13
Mais de um Participante não	2,0700	1,5226	2,8548	107,00
Região Centro Oeste	1,0390	0,7089	1,5060	3,90
Região Nordeste	1,6771	1,2407	2,2756	67,71
Região Norte	2,4041	1,5525	3,6522	140,41
Região Sudeste	1,1150	0,8616	1,4572	11,50

* Os valores da coluna Interpretação estão na forma percentual. Por exemplo: $(1,0082 - 1) * 100 = 0,8180$

Fonte: elaborado pelo autor

Pode-se observar na Tabela 9 uma variação positiva a cada unidade de renda em relação à inadimplência, o que contradiz a percepção de que, quanto maior a renda, menor a chance de inadimplência. Mutuários do sexo masculino apresentam, aproximadamente, 47,88% mais chance de inadimplência em relação ao sexo feminino. Quanto maior o financiamento maior a chance de inadimplência, conforme demonstrado pelo aumento de 3,04% a cada unidade a mais de quota de financiamento. Considerando que o mutuário comprometa em uma unidade a renda, aumenta em 9,52% a chance de ficar inadimplente. Já, em relação à variável Estado Civil, observe-se que quando considerado os mutuários Solteiros e Demais, diminui-se a chance de ficarem inadimplente quando comparado com os mutuários Casados. As duas variáveis que identificam contratos com parte da renda como informal e com participante único no financiamento imobiliário, apresentaram variação positiva de 76,13% e 107,00%, respectivamente. Em relação às regiões do Brasil, todas elas demonstram variação positiva em direção à inadimplência quando comparadas à Região Sul, tendo como extremos a Região Norte, com variação positiva de 140,41% a chance de os mutuários desta região ficarem inadimplente em relação a região Sul.

Para avaliar a capacidade preditiva do modelo iremos utilizar a curva ROC. Na Figura 3 pode-se atestar pelo gráfico gerado que o AUC foi igual a 72,40%, o que demonstra que o modelo apurado demonstrou boa capacidade preditiva.

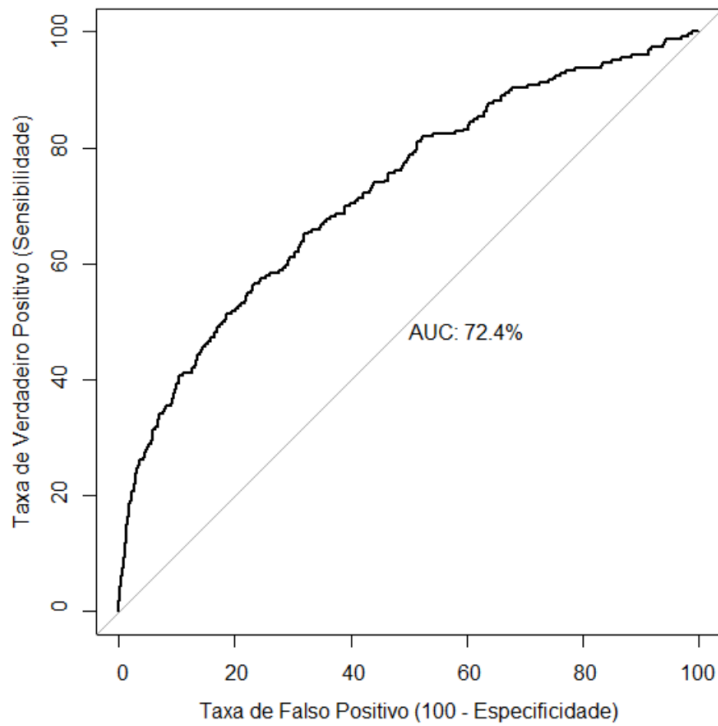


Figura 3- Curva ROC do modelo selecionado
Fonte: Elaborado pelo autor

5 Considerações Finais

O presente estudo demonstrou que a utilização do modelo de regressão logística obteve boa aderência em relação aos dados de prova, conforme ficou comprovado pela curva ROC, com área sob a curva de 72,40%.

As variáveis foram definidas com o objetivo de buscar elementos que representassem os principais elementos do financiamento, quais sejam: Mutuário, Imóvel e Dívida.

O estudo descritivo das variáveis quantitativas e qualitativas se mostrou fundamental para a definição das variáveis independentes a serem utilizadas.

Nas variáveis quantitativas foi possível avaliar que algumas variáveis demonstraram medidas muito semelhantes nos dois estratos avaliados, como pôde ser observado em relação à variável idade, na qual as medidas apresentadas ficaram praticamente iguais nos dois estratos.

Em relação às variáveis qualitativas foi feita a análise Qui-quadrado das variáveis Sexo, Estado Civil, Renda Informal, Mais de um Participante, Categoria e Estado Civil, em relação à variável dependente. Nessa análise uma variável apresentou Valor-p não significativo, acima de 0,10, Categoria do Imóvel. Esta variável foi descartada por apresentar grande valor-p, de 0,479.

Por fim chegou-se às seguintes variáveis na apuração do modelo: Renda, Sexo, Estado Civil, Quota, Comprometimento de Renda, Renda Informal e Região do Imóvel. Como pode ser atestado, os três elementos principais do financiamento se mantiveram representados, pelo mutuário: Renda, Sexo, Estado Civil, Renda Informal; pela dívida: Comprometimento de Renda, Quota e pelo Imóvel: Região.

Como o objetivo do modelo é estabelecer a probabilidade de ocorrência de inadimplência, as variáveis base para a apuração do modelo levaram em consideração, em relação às variáveis qualitativas, os valores em que se apresentaram menores índices de inadimplência, desta forma as variáveis base utilizadas no modelo foram as seguintes: Região Sul, Sexo feminino e Não possuir Renda Informal.

A apuração do modelo é importante como ferramenta de auxílio na avaliação de risco do mutuário, indicando a maior ou menor possibilidade de ocorrência da inadimplência, porém não pode ser utilizada de forma isolada, dado que muitos outros fatores interferem no momento de se avaliar a capacidade de endividamento dos proponentes ao financiamento imobiliário.

6 Referências

BOLFARINE, H.; BUSSAB, W. O. *Elementos de Amostragem*. ABE - Projeto Fisher. 1. ed. São Paulo: Edgard Blucher, 2005.

BRASIL. Lei nº 4.380 de 21 de agosto de 1964. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/l4380.htm>. Acesso em: 05 dez. 2022.

CAMARGOS, M. A. D.; CAMARGOS, M. Castro Santos; SILVA, F. Wagner; SANTOS, F. Soares dos; RODRIGUES, P. J. Fatores condicionantes de inadimplência em processos de concessão de crédito a micro e pequenas empresas do estado de Minas Gerais. *RAC-Revista de Administração Contemporânea*, 14(2), 2010.

FAWCETT, T. An Introduction to ROC Analysis. *Pattern Recognition Letters*, v 27, 2006, 861-874 p.

HOSMER, David W.; LEMESCHOW, Stanley. *Applied Logistic Regression*. 2 ed. New York: Wiley, 2000.

MENDONÇA, Mário Jorge Cardoso de. O crédito imobiliário no Brasil e sua relação com a política monetária. *Revista Brasileira de Economia*. V.67, 457-495, dez/2001.

MORAIS, Maria Piedade. *Setor habitacional, nova política habitacional*. Paraná: Finatec, Ipea, World Bank, Sedu/PR, 2001.

OHLSON, J.A. *Financial Ratios and the Probabilistic Prediction of Bankruptcy*. *Journal of Accounting Research*, 18º, 1980. 109-131 p.

SANTOS, Cláudio Hamilton M. *Políticas Federais de Habitação no Brasil: 1964/1998*. Brasília: IPEA, jul. 1999. p. 12. Disponível em: <http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=3962&Itemid=2>.