

ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

Trabalho de Formatura

**PROPOSTA DE UM MODELO ESTATÍSTICO PARA A ADMINISTRAÇÃO DA
CARTEIRA DE CRÉDITOS VENCIDOS**

AUTOR: MAURICIO SADI ANDRADE

ORIENTADORA: Profa. Dra. LINDA LEE HO

SÃO PAULO
2001

H2001
An 24/1 p

Aos meus pais.

AGRADECIMENTOS

Este trabalho somente pode se tornar realidade graças à colaboração e participação da equipe de trabalho da instituição administradora de cartões de crédito na qual baseou-se o estudo. Destaco, dentre outros, Ródnei Bernardino de Souza, Carlos Rodrigo Formigari, Ricardo Toledo de Oliveira, Gabriel Mathias Rocha, Thiago Colletti, Wanderson Bertollo e Eric Fernando. Eles, direta ou indiretamente, auxiliaram na concepção, discussão, revisão de conceitos, testes e, finalmente, resultados aqui apresentados.

Em paralelo a esta equipe, ressalto a orientação da professora Dra. Linda Lee Ho, cuja crítica e profissionalismo foram fundamentais para viabilizar a migração deste trabalho do ambiente empresarial para o acadêmico, buscando sempre o embasamento estatístico para as decisões tomadas, bem como tornando o processo de escrita mais fácil, produtivo e com encadeamento lógico.

Sou muito grato, também, ao meu irmão Luciano Sadi Andrade por todas as palavras de incentivo e pela revisão final do trabalho.

Finalmente, não poderia de ressaltar meus agradecimentos, em especial, aos meus pais Luiz Carlos de Andrade e Vera Lucia Sadi Andrade, e à minha namorada Mariana Morosini Cozza pela compreensão e paciência ao longo deste último ano de minha graduação. Seu apoio, estímulo e entusiasmo, até nos momentos mais críticos, foram fundamentais para o desenvolvimento e finalização do trabalho ora desenvolvido.

RESUMO

O presente trabalho trata do gerenciamento da carteira de inadimplentes do mercado de cartões de crédito. Reúne as informações relevantes sobre este tipo de negócio, seus principais agentes envolvidos, o fluxo financeiro das transações e o conceito de crédito rotativo. Destaca o problema da inadimplência, apontando tendências e alternativas para seu gerenciamento estratégico. Aborda a concepção, desenvolvimento, implementação e manutenção do modelo estatístico de *Charge Off Scoring*, uma ferramenta importante para a administração da carteira de créditos vencidos, a qual ocupa lugar de destaque na gestão estratégica do risco de crédito ao consumidor.

SUMÁRIO

1	INTRODUÇÃO.....	1
1.1	CONSIDERAÇÕES INICIAIS.....	1
1.2	OBJETIVO	2
1.3	APRESENTAÇÃO DOS CAPÍTULOS	3
2	CARTÕES DE CRÉDITO, INADIMPLÊNCIA E COBRANÇA.....	4
2.1	CARACTERIZAÇÃO DAS OPERAÇÕES DE CARTÃO DE CRÉDITO.....	4
2.1.1	<i>O Negócio</i>	6
2.1.2	<i>Principais Agentes.....</i>	8
2.1.3	<i>Fluxo Financeiro das Transações</i>	13
2.1.4	<i>Crédito Rotativo</i>	15
2.1.5	<i>Evolução das Operações de Cartões de Crédito no Brasil</i>	15
2.2	INADIMPLÊNCIA NAS OPERAÇÕES DE CRÉDITO AO CONSUMIDOR	19
2.2.1	<i>Inadimplência no Negócio de Cartões de Crédito</i>	24
2.3	COBRANÇA.....	27
2.3.1	<i>Formas de Cobrança</i>	28
2.3.2	<i>Provisão para Devedores Duvidosos.....</i>	31
2.3.3	<i>Resolução 2682 do Conselho Monetário Nacional.....</i>	31
3	EMBASAMENTO TEÓRICO	33
3.1	ESTATÍSTICAS DESCRITIVAS	33
3.2	ANÁLISE DE CORRELAÇÃO	34
3.3	DEFINIÇÃO DA TÉCNICA ESTATÍSTICA.....	35
3.4	REGRESSÃO LOGÍSTICA.....	37
3.4.1	<i>Análise Bivariada: Probabilidade e Odds.....</i>	37
3.4.2	<i>Modelo de Regressão Logística.....</i>	40
3.5	ADEQUAÇÃO DO MODELO	41
3.6	VARIÁVEIS DUMMY	43
4	APLICAÇÃO DAS TÉCNICAS ESTATÍSTICAS	45
4.1	CONSIDERAÇÕES INICIAIS E NECESSIDADE.....	45
4.2	O MODELO DE CHARGE OFF SCORING	47
4.2.1	<i>Definições e Conceitos.....</i>	49
4.2.2	<i>Desenho conceitual do modelo.....</i>	51
4.2.3	<i>Delineamento Amostral.....</i>	51
4.3	APLICAÇÃO DA METODOLOGIA ESTATÍSTICA	54
4.3.1	<i>Extração das Informações.....</i>	55
4.3.2	<i>Exploração das Informações.....</i>	61
4.3.3	<i>Matriz de Correlação.....</i>	64
4.3.4	<i>Aplicação da Análise Bivariada</i>	65
4.3.5	<i>Variáveis Dummy</i>	75
4.3.6	<i>Aplicação da Regressão Logística.....</i>	76
4.3.7	<i>Resultados.....</i>	81
5	ANÁLISE DOS RESULTADOS.....	83
5.1	ANÁLISE DAS VARIÁVEIS QUE COMPUSERAM O MODELO	83
5.2	PONTO DE CORTE.....	87
5.3	ANÁLISE DE DESEMPENHO	90
5.4	VALIDAÇÃO.....	93
5.5	IMPLEMENTAÇÃO.....	94

5.5.1	Estudo piloto: Definição de estratégias e resultados.....	95
5.6	MANUTENÇÃO.....	99
6	CONCLUSÃO.....	100
6.1	PRÓXIMOS PASSOS.....	101
6.2	CONSIDERAÇÕES FINAIS.....	102
7	BIBLIOGRAFIA.....	104
8	ANEXOS.....	107
8.1	ANEXO I - RESOLUÇÃO 2682.....	107
8.2	ANEXO II – VALORES CRÍTICOS DA DISTRIBUIÇÃO QUI-QUADRADO.....	113

LISTA DE FIGURAS

FIGURA 2.1	- PROCESSAMENTO DA AUTORIZAÇÃO ONLINE.....	7
FIGURA 2.2	- FLUXO FINANCEIRO DAS TRANSAÇÕES.....	14
FIGURA 2.3	- MOTIVOS DO ATRASO NOS PAGAMENTOS DE PRESTAÇÕES (EM %).....	22
FIGURA 2.4	- MOTIVOS DA PROCURA PELOS EMPRÉSTIMOS (EM %).....	24
FIGURA 2.5	- MOTIVOS DE INADIMPLÊNCIA NO CARTÃO DE CRÉDITO.....	26
FIGURA 3.1	- GRÁFICO DE $F(x) = \frac{1}{1 + e^{-\ln(\text{idade})}}$	40
FIGURA 4.1	- DESENHO CONCEITUAL DO MODELO.....	51
FIGURA 4.2	- REPRESENTAÇÃO DA COMPOSIÇÃO DA AMOSTRA.....	52
FIGURA 4.3	- LINHA DO TEMPO DA COBRANÇA DE VALORES DA VARIÁVEL MULTA/MORA.....	58
FIGURA 4.4	- MÉTODO <i>STEPWISE</i> DE SELEÇÃO DAS VARIÁVEIS.....	78
FIGURA 5.1	- DISTRIBUIÇÃO DE FREQUÊNCIA ACUMULADA DE TODA A POPULAÇÃO POR PONTUAÇÃO.....	88
FIGURA 5.2	- TESTE DE ADEQUAÇÃO DE KOLMOGOROV-SMIRNOV.....	91
FIGURA 5.3	- DISTRIBUIÇÃO DOS BONS E MAUS CLIENTES POR PONTUAÇÃO.....	92
FIGURA 5.4	- DISTRIBUIÇÃO ACUMULADA DOS BONS E MAUS CLIENTES POR PONTUAÇÃO.....	93
FIGURA 6.1	- CICLO DE MELHORIA DE PROCESSO: CONSTRUÇÃO DO MODELO.....	101
FIGURA 6.2	- CICLO CONTÍNUO DE MELHORIA: ATUALIZAÇÃO DO MODELO.....	102

LISTA DE TABELAS

TABELA 2.1	- DISTRIBUIÇÃO DOS PAGAMENTOS FINANCIADOS (EM %).....	16
TABELA 2.2	- UTILIZAÇÃO DAS MODALIDADES DE PAGAMENTO (EM %).....	17
TABELA 2.3	- EVOLUÇÃO ANUAL DE ÍNDICES DOS CARTÕES DE CRÉDITO.....	18
TABELA 2.4	- MOTIVOS DA INADIMPLÊNCIA POR TEMPO DO ATRASO(EM %).....	23
TABELA 2.5	- REGRA DE APROVISIONAMENTO.....	32
TABELA 3.1	- RELAÇÃO ENTRE AS VARIÁVEIS (ANÁLISE MULTIVARIADA).....	36
TABELA 3.2	- PROPOSTAS E COMPARAÇÃO DA DATA DE NASCIMENTO.....	38
TABELA 3.3	- TABELA DE CONTINGÊNCIA.....	39
TABELA 3.4	- EXEMPLO DE CODIFICAÇÃO DE VARIÁVEIS QUALITATIVAS.....	44
TABELA 4.1	- COMPOSIÇÃO DAS BASES DE DADOS AMOSTRAIS.....	53
TABELA 4.2	- SIGLAS MNEMÔNICAS DAS VARIÁVEIS.....	62
TABELA 4.3	- ESTATÍSTICAS DESCRITIVAS DA BASE DE DADOS DE DESENVOLVIMENTO.....	63
TABELA 4.4	- COMPOSIÇÃO DEFINITIVA DAS BASES DE DADOS AMOSTRAIS.....	63
TABELA 4.5	- MATRIZ DE CORRELAÇÃO DAS VARIÁVEIS CONTÍNUAS.....	64
TABELA 4.6	- ANÁLISE BIVARIADA DO RISCO DE INADIMPLIR EM T x VARIÁVEL RESPOSTA.....	66
TABELA 4.7	- VALORES OBSERVADOS E ESPERADOS PARA <i>BEHAVIOR SCORE B</i> E <i>C</i>	67
TABELA 4.8	- VALORES OBSERVADOS E ESPERADOS PARA OS GRUPOS DE <i>BEHAVIOR SCORE</i>	68

TABELA 4.9 – ANÁLISE BIVARIADA DO RISCO DE INADIMPLIR EM T-1 X VARIÁVEL RESPOSTA.....	69
TABELA 4.10 – ANÁLISE BIVARIADA DO PRODUTO X VARIÁVEL RESPOSTA.....	70
TABELA 4.11 – ANÁLISE BIVARIADA DA BANDEIRA X VARIÁVEL RESPOSTA	71
TABELA 4.12 – ANÁLISE BIVARIADA DO LOCAL DE RESIDÊNCIA X VARIÁVEL RESPOSTA.....	73
TABELA 4.13 – CONVERSÃO DAS VARIÁVEIS INDEPENDENTES DISCRETAS EM VARIÁVEIS <i>DUMMY</i>	75
TABELA 4.14 – VARIÁVEIS FINAIS PARA ENTRADA NO MODELO.....	76
TABELA 4.15 – SUMÁRIO DO MÉTODO DE SELEÇÃO <i>STEPWISE</i>	79
TABELA 4.16 – ANÁLISE DOS COEFICIENTES DE MÁXIMA VEROSSIMILHANÇA.....	80
TABELA 4.17 – TABELA DE PARÂMETROS (RESULTADOS DO MODELO)	82
TABELA 5.1 – FAIXAS DE PONTUAÇÃO DE “BONS” E “MAUS”	89
TABELA 5.2 – CLASSIFICAÇÃO DAS PONTUAÇÕES (x 100) DO MODELO	89
TABELA 5.3 – DISTRIBUIÇÃO DOS “BONS” E “MAUS” POR ESCORE	94
TABELA 5.4 – ESTRATÉGIAS ADOTADAS PARA O ESTUDO PILOTO.....	96
TABELA 5.5 – RESULTADO DAS CONTAS ESCORADAS APÓS IMPLEMENTAÇÃO DAS ESTRATÉGIAS.....	97

1 INTRODUÇÃO

1.1 *Considerações Iniciais*

O estágio realizado na administradora de cartões de crédito de um banco de varejo nacional, mais especificamente em sua área de recuperação de crédito, despertou tanto o tema como a motivação para que este trabalho fosse realizado. Diante da busca pelo incremento na produtividade e pelo maior foco do trabalho, identificou-se a necessidade de criação de uma ferramenta que agilizasse o processo de cobrança, reduzindo consequentemente seu custo.

A instituição em questão já emprega alguns modelos de gerenciamento de sua carteira: modelos de análise de crédito, de previsão de comportamento e de recuperação de crédito, porém este último foi concebido apenas para clientes que apresentam poucos dias em atraso de sua fatura mensal (menos de 80 dias), não havendo qualquer tipo de classificação dos associados com maiores atrasos, dificultando o direcionamento e o foco da intervenção a ser feita sobre estes.

Com a promulgação da Resolução 2682 do Banco Central, a necessidade de instauração deste novo modelo aumentou, uma vez que esta nova resolução passou a regulamentar com maior severidade a Provisão para Devedores Duvidosos, foco principal de atuação da Área de Recuperação de Crédito.

Assim sendo, objetivando-se construir uma ferramenta que pudesse solucionar o problema mencionado anteriormente, bem como aumentar a produtividade e melhorar a abordagem ao cliente inadimplente utilizando padrões diferenciados, decidiu-se pela confecção de um modelo para recuperação do crédito de inadimplentes com muitos dias em atraso, ou, conforme denominaremos Modelo de *Charge-Off Scoring*.

1.2 Objetivo

Este trabalho tem como objetivo a elaboração de um modelo, com base em técnicas estatísticas, que permite diferenciar os associados inadimplentes com alta propensão a financiar sua dívida, daqueles que notadamente não têm condições de zerar o seu saldo no atual momento. Esta intenção se baseia na necessidade de direcionamento dos esforços intervencionistas para a recuperação do crédito concedido por uma administradora de cartões de crédito, utilizando informações de fatores quantitativos e também qualitativos.

Já existem, é verdade, na empresa em questão, alguns outros modelos que suprem parcialmente a necessidade descrita no parágrafo anterior, cobrindo a parcela de associados que estão em inadimplência há pouco menos de 3 meses. Estes modelos são direcionados a identificar aqueles clientes que têm propensão a quitar suas dívidas, separando-os daqueles que não o farão no momento por algum motivo. E são justamente estes que comporão a base de estudo para a confecção do modelo proposto neste trabalho.

Os clientes que ultrapassarem 3 meses em inadimplência serão o foco do estudo, estudo este que procurará, também, obter uma variável resposta diferente da anterior: a propensão do associado a começar a financiar sua dívida.

1.3 Apresentação dos Capítulos

A seguinte estrutura foi empregada no desenvolvimento deste trabalho:

- **Primeiro capítulo – Introdução**

Apresenta uma breve descrição do trabalho, expondo o objetivo e a motivação para a elaboração deste.

- **Segundo capítulo – Cartão de Crédito, Inadimplência e Cobrança**

Caracteriza o mercado de cartão de crédito, descrevendo o negócio, seus principais agentes, suas operações e a evolução histórica recente das operações de cartão de crédito no país. Contém, também, um panorama sobre a inadimplência e a recuperação dos saldos dos devedores, destacando as formas de cobrança empregadas, a composição da Provisão para Devedores Duvidosos com base na Resolução 2682 do Banco Central.

- **Terceiro capítulo – Embasamento Teórico**

Exibe todas as técnicas estatísticas para modelagem de dados aplicadas no estudo, assim como o motivo da escolha destas.

- **Quarto capítulo – Aplicação das Técnicas Estatísticas**

Faz uma descrição da necessidade de confecção do modelo, apresentando em seguida, a utilização das técnicas exibidas no Capítulo 3, desde a seleção da base de dados até o modelo final obtido.

- **Quinto capítulo – Análise dos Resultados**

Discute os resultados encontrados e a efetividade do modelo, passando por sua validação, implementação (com um estudo piloto sobre as estratégias adotadas com base nos resultados do modelo) e manutenção.

- **Sexto capítulo – Conclusão**

Expõe um resumo dos resultados alcançados, os próximos passos para o modelo e tudo que pode ser aprendido durante sua construção.

- **Sétimo capítulo – Bibliografia**

- **Oitavo capítulo – Anexos**

2 CARTÕES DE CRÉDITO, INADIMPLÊNCIA E COBRANÇA

Com a finalidade de ambientar o leitor diante de todo o mercado de cartões de crédito, com suas particularidades e detalhes específicos, serão expostas, neste capítulo, as características do setor de cartões de crédito, abrangendo dentre outros itens o relacionamento entre os principais agentes e o fluxo financeiro deste negócio.

Em seguida, será feita uma análise do problema da inadimplência no mercado de concessão de crédito, destacando em particular, o negócio de cartões de crédito e a administradora na qual o modelo estatístico foi desenvolvido.

Apresentaremos, então, uma breve explanação sobre a provisão para devedores duvidosos, mais comumente conhecida como PDD, sua função, detalhes e regras do Banco Central para sua constituição. Serão descritas, também, as principais formas e estratégias empregadas na cobrança e recuperação do crédito em atraso, apresentando como estão sendo utilizados os modelos de predição de inadimplência, e evidenciando a necessidade de utilização do modelo de *charge-off scoring* como uma ferramenta estratégica de fundamental importância para o gerenciamento da carteira de inadimplência.

Veremos, a seguir, detalhadamente cada um deles, ressaltando suas inter-relações e interdependências.

2.1 *Caracterização das Operações de Cartão de Crédito*

“Emprestar dinheiro é uma das profissões mais antigas do mundo. O Código de Hamurabi, surgido em 1800 a.C., apresentava diversas seções relativas à regulamentação do crédito na Babilônia” (Caouette, Altman & Narayanam, 1999, p.1).

“A palavra ‘crédito’, dependendo da situação, pode ter vários significados. Para a atividade bancária, crédito é definido como a entrega de um valor presente mediante uma promessa de pagamento” (Silva, 1997, p.63). Como se pode constatar, a concessão

de crédito significa confiar que outrem irá honrar seus compromissos, nas datas pactuadas anteriormente.

O crédito é um elemento presente em praticamente todas as políticas financeiras das empresas comerciais e industriais, como forma de alavancar as vendas ou suprir eventuais necessidades de caixa, realização de aquisições ou aumento da capacidade de produção. Diante destas necessidades, a função dos bancos é promover a intermediação financeira, ou seja, administrar recursos de terceiros e fornecer crédito ao comércio, à indústria, às empresas prestadoras de serviços e às pessoas físicas.

Com a estabilidade econômica, proporcionada pelo Plano Real, os bancos, privados dos benefícios do floating (tempo decorrido entre o envio de recursos pelo devedor e o momento em que os recursos estão disponíveis para o credor), acirraram a competitividade e perceberam a necessidade de investir pesadamente em mecanismos, para que seus negócios sejam efetuados com maior agilidade, segurança e eficiência, possibilitando redução de custos e aumento da lucratividade.

A tendência de queda nas taxas de juros e a estabilidade do câmbio indicam que o crédito será a grande fonte de recursos para os bancos nos próximos anos, obrigando o setor bancário a cumprir com maior ênfase sua função básica de fomento à economia.

O negócio de cartões de crédito é, em virtude da sua estrutura operacional, mais arriscado do que a maioria dos outros empréstimos ao consumidor – e as perdas relacionadas a ele são mais altas (detalhes no Seção 2.2.1). Entretanto, se o risco for previsto no estabelecimento de taxas, e elas não afetarem negativamente o mercado, então não será causa para maiores preocupações. Da mesma forma, é mais arriscado ter negócios em alguns países ou estados do que em outros. Assim sendo, administrar esse risco é, portanto, uma das principais responsabilidades da gestão estratégica do crédito ao consumidor.

2.1.1 O Negócio

Os cartões de crédito são utilizados na aquisição de bens e serviços em estabelecimentos que estejam credenciados por empresas adquirentes (Seção 2.1.2), com a participação de bandeiras e de emissores (ou administradoras) de cartões. Eles servem como meio de pagamento de compras em uma rede de estabelecimentos conveniada às bandeiras, além de permitirem saques em caixas eletrônicos no país e no exterior, funcionando como crédito automático. Sua venda é feita por emissores, filiados a bandeiras internacionais, que podem ter ou não uma administradora própria. Dentre os principais emissores podemos destacar a Credicard, a American Express e os bancos nacionais e estrangeiros.

Não há cobrança de juros sobre as compras efetuadas no período compreendido entre o dia da transação e a data de vencimento da fatura mensal. Os juros somente passam a incidir no caso do uso do crédito rotativo (Seção 2.1.4), ou quando o portador utiliza o cartão para efetuar saques em caixas eletrônicos. Os cartões magnéticos (comumente conhecidos como plásticos) têm, em geral, validade de dois anos, quando então são renovados pelas administradoras.

Podemos dividir os tipos de produtos das administradoras de cartões de crédito em: cartões domésticos, internacionais, múltiplos, empresariais, de afinidades e *co-brandeds*. Os primeiros podem ser utilizados exclusivamente no país, enquanto que os internacionais possuem também a aceitação no exterior. Os cartões múltiplos agregam as funções de crédito e de débito, ou seja, além das formas de utilização referentes aos cartões de crédito já descritas, também desempenham funções bancárias como movimentação de contas correntes, saques e extratos.

Os cartões de débito são, em geral, emitidos por bancos e usados para saques, pagamentos de contas e movimentar uma conta corrente. Podem existir, no verso destes cartões, logomarcas, tais como o Cheque Eletrônico da TecBan e o Rede Shop da Credicard, os quais permitem ao usuário adquirir bens ou serviços em estabelecimentos comerciais. O valor da transação é debitado automaticamente da conta do portador do

cartão. As bandeiras internacionais Visa e Mastercard também possuem marcas (Visa Electron, por exemplo) que também permitem, com o uso dos cartões de débito, o débito do valor das compras em conta corrente.

Voltando aos cartões de crédito, a autorização de suas transações envolve uma rede de telecomunicações entre as bandeiras, emissores, adquirentes e processadoras independentes. Uma autorização com cartão de crédito leva, no máximo 10 segundos, enquanto que com o cartão de débito não se estende por mais de 4 segundos em média. O fluxo que envolve a autorização das transações pode ser visualizado através da Figura 2.1 a seguir:

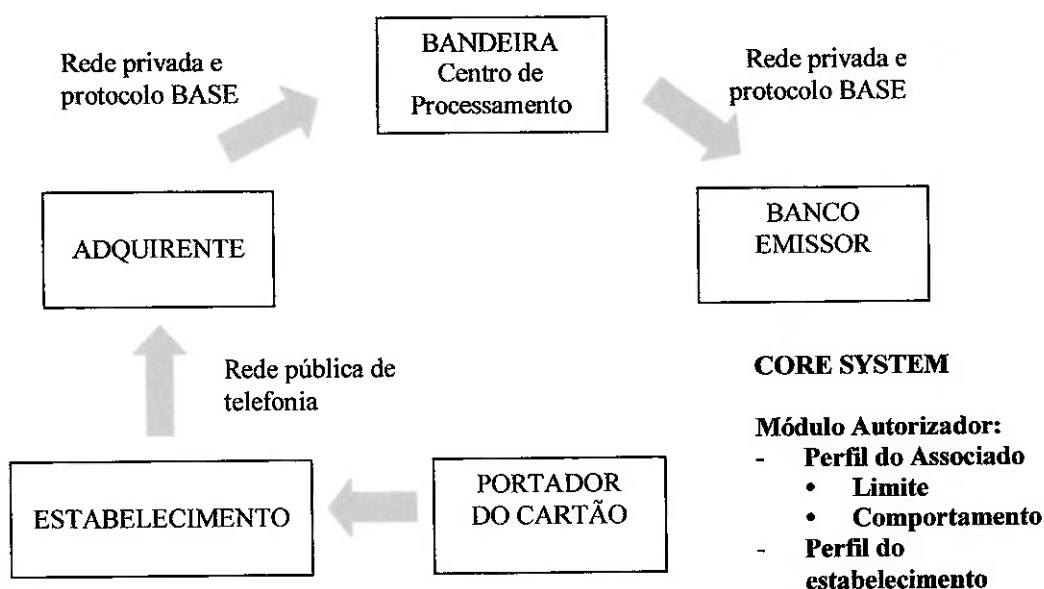


Figura 2.1 - Processamento da Autorização Online – Adaptada de Souza (2000)

Existe, atualmente, um movimento de internacionalização dos cartões de crédito, ou seja, há a tendência dos administradores em substituir no mercado os cartões domésticos pelos internacionais. Da mesma maneira, existe uma mudança em direção aos cartões múltiplos. Os grandes bancos presentes no mercado nacional como o Real, o Unibanco, o Santander aderiram aos cartões múltiplos, assim como o Bradesco, por

exemplo, que até o final de 2001 incluirá em sua base de cartões as opções de crédito e débito no mesmo produto.

2.1.2 Principais Agentes

O setor de cartões de crédito é constituído por diversos agentes que se inter-relacionam de modo a permitir que as transações de crédito aconteçam. No Brasil, os principais agentes são representados pelas administradoras, também chamadas de emissores, que além de administrar podem processar cartões, bandeiras, empresas adquirentes e processadoras independentes.

Os administradores (ou emissores)

Os administradores são praticamente os donos dos cartões de crédito. Realizam serviços de administração e podem também optar por fazer o processamento dos cartões (serviço que pode ser deixado a cargo de processadoras independentes). Sucintamente, emitem em favor de uma pessoa um cartão de crédito, pessoal e intransferível, que lhe permite realizar compras em estabelecimentos comerciais filiados à rede dos adquirentes (que serão apresentados mais adiante, ainda nesta Seção). Por intermédio destes adquirentes, pagam a estes estabelecimentos após 30 dias em geral, e, descontam dos mesmos uma taxa sobre o valor da compra, que pode variar de 2 a 5%. No caso das compras internacionais, o prazo para pagamento dos estabelecimentos é de 2 dias. Do portador do cartão é cobrada uma taxa de anuidade que pode oscilar entre US\$15,00 e US\$90,00 e taxas de juros incidentes sobre o financiamento das transações, que podem variar entre 4 a 13% ao mês.

São mantidos contratos com os portadores de cartões, com os estabelecimentos comerciais filiados e com uma instituição bancária que financia as compras realizadas por meio deste cartão. É a administradora quem paga aos estabelecimentos comerciais e arca com os riscos em caso de inadimplência dos associados. O pagamento pelo titular

do cartão pode ser feito em até 30 dias após a compra, sem a incidência de juros. Se o portador optar por utilizar o crédito rotativo, irá saldar uma parte do débito, em geral 20%, e o restante será financiado.

Criadas por bancos ou a eles associadas, as empresas administradoras de cartões de crédito obtêm recursos para financiamento das compras com cartões por meio de transações financeiras realizadas junto a estes bancos. Em meados de 2001, os nove maiores emissores de cartões de crédito no Brasil eram: Credicard, Itaú, Banco do Brasil, Bradesco, Unibanco, ABN Amro – Real, Santander, Fininvest e American Express.

As bandeiras

Atuam no país hoje somente quatro bandeiras: Visa, Mastercard, American Express e Diners.

As bandeiras, a exemplo da Visa e da Mastercard, foram criadas nos Estados Unidos a partir de associações de bancos com o objetivo de emitir cartões de crédito. Reúnem milhares de instituições financeiras em todo o mundo. O Diners Club, que surgiu nos Estados Unidos na década de 50, foi comprado da Continental Insurance Corporation pelo Citigroup Inc. em 1981.

A American Express é uma empresa norte-americana que está em muitos países por meio de filiais. As bandeiras Visa, Mastercard e Diners operam por meio de sistemas abertos de franquias, que não é o caso da American Express. No Brasil, o Diners Club é uma franquia da Credicard e a American Express opera por meio de uma filial instalada no país.

O sistema aberto permite às instituições financeiras associadas, por exemplo, emitir cartões de crédito próprios usando a marca das bandeiras e tornarem-se responsáveis pela administração destes cartões. Isso envolve decisões sobre o gerenciamento e processamento do produto. Tais atividades podem ser terceirizadas pelos bancos.

O caso da American Express é diferente. Além de bandeira, a empresa é também uma administradora, emissora de cartões e responsável pelo credenciamento dos estabelecimentos comerciais para aceitação de seus produtos. Acompanhando uma estratégia adotada no exterior, fez acordos com bancos e empresas permitindo que estes emitissem seus cartões de crédito. No Brasil, a bandeira tem acordos com o HSBC Bamerindus, Bilbao Vizcaya, BCN e Sony do Brasil.

A marca das bandeiras está estampada nos cartões e identifica a rede de aceitação em estabelecimentos filiados. As bandeiras desenvolvem novos produtos e serviços, criam programas de marketing, publicidade e promoções; têm uma linha de produtos e serviços tanto para os cartões de crédito, quanto para os de débito, assim como garantem o intercâmbio das informações decorrentes dos pagamentos feitos com seus cartões permitindo que as mesmas cheguem ao seu emissor através de redes adquirentes. Se relacionam, em geral, com as administradoras ou com as processadoras independentes, com os adquirentes e com os fabricantes de cartões, fazendo também sua homologação, para que estes possam operar com os cartões de crédito e débito que possuem sua marca.

Os bancos membros pagam várias taxas para as bandeiras, entre elas *royalties* por volume de cartões e transações e uma taxa mensal pelo direito do uso das marcas no cartão. No caso da Mastercard, é cobrada uma taxa dos emissores de cartões de crédito por transação que trafega na sua rede de cerca de US\$0,0235 e também uma taxa que gira em torno de US\$800,00 por mês, referentes a serviços de comunicação prestados aos mesmos, entre os quais linhas de comunicação e conexão da sua central de processamento de dados com os emissores. É cobrada, também, uma taxa anual dos fabricantes de plásticos de aproximadamente US\$30 mil.

Os investimentos das bandeiras concentram-se principalmente em marketing para divulgação da sua marca e desenvolvimento de novos produtos, serviços e tecnologia.

No portfólio de produtos das bandeiras encontram-se os cartões de crédito nacionais, internacionais, empresarias, *co-branded*, de afinidades, cheques de viagem (*travellers cheques*), cartões de débito e cartões com *chip*. A bandeira oferece vários serviços agregados aos cartões no país e no exterior que podem ser terceirizados. Nesse

caso, as bandeiras pagam um valor mensal para as empresas prestadoras destes serviços por número de cartões, por tipo de plástico e de serviço oferecido.

As empresas adquirentes

As empresas adquirentes (também chamadas de *acquirers*) são responsáveis pelo credenciamento dos estabelecimentos comerciais e pela intermediação dos fluxos financeiros entre os emissores e os estabelecimentos que aceitam o produto.

No Brasil, existem 3 adquirentes: Visanet, que credencia os estabelecimentos que aceitam as bandeiras de crédito e débito Visa (crédito) e Visa Electron (débito); Redecard, responsável pelos estabelecimentos que trabalham com as bandeiras Mastercard (crédito), Maestro (débito), Diners e Rede Shop (que pertence à Credicard); e American Express, que credencia diretamente os estabelecimentos. A Visanet e a Redecard têm participação das bandeiras e de instituições financeiras.

Com a criação dos adquirentes Visanet e Redecard em 1996, o relacionamento entre os emissores e os estabelecimentos comerciais foi muito facilitado. Até então, a responsabilidade pelo credenciamento de estabelecimentos e pelas transações realizadas pelos cartões Visa, Mastercard e Diners era dos próprios emissores dos cartões. O Bradesco, a Credicard e o Banco do Brasil, por exemplo, eram quem fazia o credenciamento dos estabelecimentos comerciais. O surgimento dos adquirentes foi motivado pelo crescimento do negócio de cartões de crédito após o Plano Real, e pela possibilidade dos emissores trabalharem com as bandeiras Visa e Mastercard simultaneamente a partir de 1996. Até aquele momento, a Credicard, por exemplo, credenciava os estabelecimentos comerciais que aceitavam a bandeira Mastercard e Diners, e Bradesco e Banco do Brasil, entre outros, os da bandeira Visa. Com os emissores operando com mais de um bandeira, o mercado procurou uma forma de concentrar as transações, facilitando o relacionamento entre administradoras e estabelecimentos.

As redes adquirentes fazem a captura dos dados e enviam as informações das transações efetuadas com os cartões de crédito e débito aos emissores, que concedem a

autorização para que a venda ou pagamento de um serviço seja concluído. São, também, responsáveis pelos pagamentos aos lojistas das compras efetuadas com os cartões.

Os adquirentes são também responsáveis por equipar as lojas com máquinas, o POS (*point of sale*) ou PDV (ponto de venda), este último usado principalmente em supermercados, além de fornecerem bobinas para as máquinas POS, que imprimem os comprovantes de compra. Estas empresas também fazem a sinalização do estabelecimento com a colocação de adesivos nas vitrines, indicando aos usuários quais os cartões são aceitos naquele local.

As processadoras independentes

As processadoras independentes são assim chamadas por não pertencerem aos bancos. Realizam as atividades operacionais envolvidas na gestão completa de um cartão em nome de um emissor. Atuam junto aos bancos e empresas oferecendo uma gama de serviços que vai desde a administração, processamento de cartões e envio de faturas aos associados até o atendimento aos clientes, passando pela concessão de crédito, prevenção à fraude e cobrança.

A CSU CardSystem foi a pioneira no Brasil. Além dela, existem mais duas empresas que concorrem no negócio de processamento de cartões: Equifax/Unnisa e EDS.

A Credicard, que realizou uma reestruturação em maio de 1999, tem uma empresa processadora de cartões que presta serviços a terceiros. Esta processadora, entretanto, não pode ser considerada independente, uma vez que a administradora é constituída por três bancos (Citibank, Itaú e Unibanco), diferentemente das outras três citadas anteriormente.

As processadoras cobram uma taxa por cartão administrado que, em geral, pode variar de US\$20,00 a US\$50,00 por cartão ao ano. A taxa envolve serviços como: análise de crédito, cadastramento de futuro usuário do cartão no sistema da empresa, gravação do nome do portador e da trilha magnética, postagem e envio do cartão por Sedex especial. O preço cobrado dos emissores também inclui o cálculo de juros em

caso de atraso do pagamento, envio da fatura mensal para o cliente, atendimento telefônico (*call center*), controle de fraudes e inadimplência.

O surgimento destas processadoras independentes veio estimular as instituições financeiras e empresas do segmento de varejo, como supermercados e lojas de departamentos, a desenvolverem seus próprios cartões.

2.1.3 Fluxo Financeiro das Transações

O fluxo financeiro das transações se inicia com a utilização do cartão por parte do portador como forma de pagamento na compra de um bem ou serviço. As informações da transação (valor) e do cartão de crédito (número) seguem do comerciante para o adquirente. Estas mesmas informações são redirecionadas pelo adquirente para o administrador (ou emissor) do cartão.

Uma vez alcançando o administrador, os dados da transação são por este processados de modo a verificar se existe a possibilidade de autorização da compra ou não.

Tanto em caso negativo quanto em caso positivo, a resposta retorna do emissor para o adquirente, deste para o estabelecimento e finalmente chega ao portador do cartão de crédito. Neste momento ocorre uma diferenciação para as duas possibilidades acima descritas: sendo a resposta negativa, o estabelecimento informa ao comprador que este não poderá efetuar o pagamento através do cartão de crédito, propondo-lhe outras formas de se alcançar o mesmo fim, e, caso a resposta seja positiva, um novo fluxo se inicia, o qual trata da troca de dados referente ao processo de pagamento.

Assim, é emitido por parte do vendedor o pedido de reembolso do valor da transação; este pedido passa pelo adquirente chegando até o emissor. Este, por sua vez, efetua o pagamento do valor devido ao adquirente 27 dias após o processamento da compra com um desconto (conhecido como taxa de intercâmbio) de 1,00% se a transação for eletrônica e 1,44% se for manual. Para as transações internacionais, o repasse do emissor ao adquirente acontece 2 dias após o processamento.

A partir daí, o pagamento é redirecionado pelo adquirente ao estabelecimento comercial, em média, depois de 30 dias, retendo consigo por volta de 5%. Fechando este ciclo, a administradora emite uma fatura mensal, cobrando de seu associado o valor financeiro referente a esta transação de compra, que, em média efetua o pagamento 25 dias após a compra do bem ou do serviço.

A Figura 2.2 ilustra claramente este fluxo financeiro das transações.

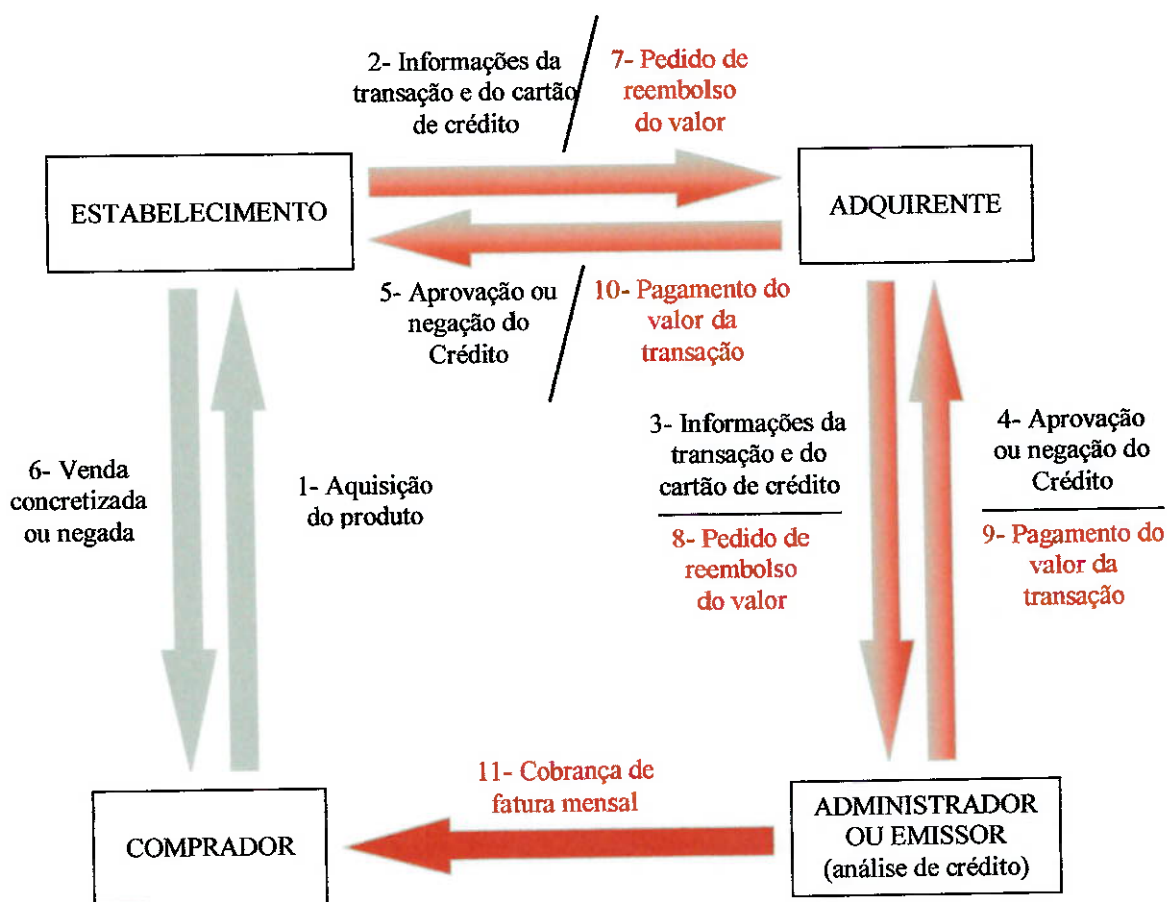


Figura 2.2 - Fluxo financeiro das transações – Adaptada de Souza (2000)

2.1.4 *Crédito Rotativo*

Uma das mais importantes fontes de receita para as administradoras de cartões de crédito é o financiamento. No Brasil, entre 40 e 45% do total dos gastos efetuados com cartões são financiados a partir do vencimento, incidindo sobre estes uma taxa de juros, em média de 11% ao mês.

Financiar no cartão de crédito significa pagar um valor entre o mínimo permitido pelas administradoras e o valor total da dívida que consta da fatura do cliente. A média do mercado para o valor mínimo a ser pago é de 20% em relação ao total da fatura. Ao pagar o valor mínimo, o cliente está optando automaticamente pelo crédito rotativo, ou seja, está financiando suas compras pelo cartão de crédito.

Para as despesas internacionais, entretanto, por determinação do Banco Central, não é permitido o financiamento através do cartão de crédito; estas devem, portanto, serem pagas integralmente.

A propensão a financiar no Brasil varia de 40 a 45%, contra 80% nos Estados Unidos; um dos fatores que podem explicar esta grande diferença é a distância entre as taxas de juros comumente aplicadas nos dois países. O crédito rotativo correspondia, em 1999, a 40% da receita total obtida com o negócio de cartões de crédito. Algumas administradoras revelaram que esta se trata de sua principal fonte de receita, seguida pela anuidade.

O não pagamento de pelo menos 15% da fatura até sua data de vencimento significa, para a administradora, que o associado encontra-se em atraso com seu pagamento e, nesse caso, diz-se que ele está inadimplente (a discussão sobre inadimplência será melhor explorada adiante, na Seção 2.2.1).

2.1.5 *Evolução das Operações de Cartões de Crédito no Brasil*

O uso mais freqüente dos cartões de crédito é recente no país. Segundo dados do Panorama Setorial da Gazeta Mercantil (1999), em 1996 o cheque pré-datado foi

responsável por 35% dos pagamentos financiados, enquanto que o cartão de crédito respondeu por 24%. No final de 1998, entretanto, essa participação inverteu-se: cheque pré-datado caiu para 26% e o cartão de crédito alcançou 34,7%, conforme Tabela 2.1.

Financiamento	1996	1997	1998
Cartão de Crédito	24,0	22,3	34,7
Financiamento de bens duráveis	35,6	41,7	29,5
Cheque pré-datado	34,9	27,1	26,0
Cartão de Lojas	5,5	8,9	9,8

Tabela 2.1 – Distribuição dos pagamentos financiados (em %) – Extraída do Panorama Setorial da Gazeta Mercantil (1999)¹

Da mesma forma, o uso do cartão de crédito como modalidade de pagamento (não necessariamente como fonte de crédito ao consumidor) também tem apresentado um crescimento significativo. Em 1996, ele representava 5,8% do total das modalidades de pagamento utilizadas e, em 1998, essa participação já estava em 8%. A Tabela 2.2 mostra que o cartão de crédito ganhou participação do mercado do cheque, que teve sua participação reduzida de 21,2% (à vista mais pré-datado) para 16,5%, e do financiamento, que foi de 8,7% para 6,6%. Os demais meios de pagamento também tiveram um aumento em seus percentuais.

¹Base: População economicamente ativa urbana com mais de 16 anos e renda individual acima de R\$300,00.

Financiamento	1996	1997	1998
Em dinheiro	62,2	64,6	64,9
Cheque à vista	11,8	9,9	9,9
Cartão de Crédito	5,8	4,8	8,0
Cheque pré-datado	9,4	6,9	6,6
Financiamento de bens duráveis	8,7	10,4	6,6
Cartão de Lojas	1,4	1,9	2,1
Cartão de débito	0,7	1,6	2,1

Tabela 2.2 – Utilização das modalidades de pagamento (em %) – Extraída do Panorama Setorial da Gazeta Mercantil (1999)²

Esses dois resultados , aumento nas participações dos pagamentos financiados e dos meios de pagamento, mostram uma tendência de substituição dos cheques pelo dinheiro “de plástico” no varejo; esta tendência pode também ser reforçada pelo acentuado crescimento dos números do setor, apresentados na Tabela 2.3.

²Base: População economicamente ativa urbana com mais de 16 anos e renda individual acima de R\$300,00.

Período	Número de Cartões (Milhões)	VARIAÇÃO (+/-)	Nº. de Transações (Milhões)	VARIAÇÃO (+/-)	Valor de Transações (US\$) (Bilhões)	VARIAÇÃO (+/-)
1991	7,9	-	105,7	-	5,2	-
1992	7,8	(1,27)	151,6	43,42	5,1	(1,92)
1993	8,4	7,69	199,9	31,86	6,3	23,53
1994	11,2	33,33	210,3	5,20	10,3	63,49
1995	14,3	27,68	319	51,69	21,3	106,80
1996	17,2	20,28	437,1	37,02	25,5	19,72
1997	19,3	12,21	516,7	18,21	27,8	9,02
1998	22	13,99	641,2	24,10	32	15,11
1999	23,6	7,27	770	20,09	41,6	30,00
2000	27,95	18,43	1.038	29,92	50,36	21,06

Tabela 2.3 – Evolução anual de Índices dos Cartões de Crédito.

Fonte: <<http://www.abecs.org.br>>

De acordo com os dados da Federação do Comércio do Estado de São Paulo (FCESP), em agosto de 2001, a participação do cartão de crédito no faturamento do varejo, voltou a cair. O cartão estava com 14,6% do faturamento em abril de 2000 e, a partir dessa data não parou mais de subir: 17,9% em agosto de 2000, 20,9% em novembro, 27,7% em abril deste ano e, agora, em agosto, a surpresa, 17,1%. A expectativa era que o cartão de crédito estivesse conquistando fatias cada vez maiores de clientes, uma tendência que deveria acompanhar a modernização e informatização da economia. Os segmentos ganhadores foram o faturamento à vista, de 41,7%, em abril deste ano, para 45,8%, em agosto, e o cheque pré-datado, de 26,5% para 32,1%. A modalidade à vista continua sendo a mais importante em todos os segmentos, incluindo duráveis que deveria ter uma participação maior de financiamentos na composição das vendas, em tese, por ser composto de bens de valor unitário mais alto.

2.2 *Inadimplência nas Operações de Crédito ao Consumidor*

De acordo com os dados publicados pelo Banco Central, houve um crescimento, pós Plano Real, do sistema de crédito ao consumidor no Brasil, superando os R\$20 bilhões de saldo no primeiro semestre de 1997, o equivalente a 35% sobre o saldo de 1994 (R\$15 bilhões). Durante o período desse crescimento, e com o objetivo de controlá-lo, o governo passou a adotar medidas restritivas que pudessem adequar o ritmo de crescimento da demanda ao programa de estabilização depois da implantação do Plano Real.

O aumento do crédito, aliado às medidas restritivas, atingiram de tal modo o consumidor, que levou o país à redescoberta da inadimplência nas operações de empréstimos que, antes, eram mascaradas pela alta inflação e salários indexados.

Os anos de 95, 96 e 97 foram marcados por um aumento considerável da inadimplência, em virtude, principalmente, do aumento da taxa de juros, fazendo com que o número de títulos protestados em 97, por exemplo, fosse 13% superior ao ano de 96 e a participação de pessoas físicas nesse índice subisse de 23% para 28%, segundo dados da Associação das Empresas de Recuperação de Crédito (ASERC) 2000.

Na realidade, o país não estava preparado para receber o aumento do crédito ao consumidor e este não havia se habituado à estabilidade econômica e à necessidade de conscientização quanto ao controle de suas contas. As pessoas tiveram seu poder aquisitivo acrescido com a volta dos financiamentos, porém nem todos souberam utilizar esta vantagem adequadamente às suas condições financeiras. Como resultado, a inadimplência acabou explodindo e a maior parte das instituições não estava preparada para isto.

O primeiro mecanismo de proteção contra o aumento da inadimplência adotado pelas instituições financeiras foi a elevação generalizada do *spread* (diferença entre as taxas de captação e de empréstimos aos consumidores adotadas pelos bancos), com o conseqüente aumento das taxas de juros. Esse aumento tornou os empréstimos mais onerosos e difíceis de serem liquidados, contribuindo, assim, para o crescimento da

inadimplência. Um estudo do Banco Central (1999) revelou que 35% do *spread* dos bancos era proveniente da inadimplência no Brasil.

Obviamente, um dos fatores que contribuiu para a elevação do *spread* bancário, e consequentemente a inadimplência no país, foi o aumento e a volatilidade da taxa básica de juros (Selic) nesses últimos anos, resultado da maior volatilidade do mercado financeiro internacional. Entretanto, com a queda e a estabilidade da taxa Selic, as quedas nos juros praticados pelos bancos foram inferiores às expectativas do Banco Central.

A ineficiência econômica – aumento da inadimplência “protegida” pelo aumento da taxa de juros, que provoca mais inadimplência – gerada nesses anos posteriores ao Plano Real e, particularmente, nos três últimos anos, fizeram as instituições financeiras no Brasil repensar os seus mecanismos de concessão, manutenção e recuperação de créditos. Muitas discussões entre acadêmicos, reguladores e profissionais bancários têm demonstrado que a principal causa das perdas de crédito associadas a essa ineficiência econômica é “a falta de uma gestão mais profissional, e até mesmo estratégica, do risco de crédito, que possa melhor identificar, mensurar e administrar os vários tipos de riscos intrínsecos – risco do cliente (probabilidade de que o cliente torne-se inadimplente em um horizonte específico) do tipo de operação (do cartão de crédito, por exemplo) e da concentração (em vez da diversificação) – à atividade de crédito” (Douat e Sicsú, 1999).

Neste sentido, a gestão do risco de crédito é o grande desafio dos próximos anos no Brasil. Para isto, as instituições financeiras nacionais começam a adotar a “administração estratégica do risco de crédito” (Douat e Sicsú, 1999) com a utilização de novos conceitos e instrumentos que dêem sustentação a essa nova abordagem.

Dentro dessa perspectiva, uma das soluções propostas por Malaga (1999) é a manutenção de baixas taxas de juros com a oferta de crédito para tomadores selecionados. Ele afirma que a estratégia de compensar a perda de crédito com o aumento da taxa de juros é perigosa e, no Brasil, essa prática virou regra.

Entretanto, para evitar a seleção adversa – em virtude da assimetria de informação, entre tomador e credor, sobre a probabilidade de inadimplência do tomador – Malaga (1999) sugere que as instituições credoras forcem uma sinalização de baixo

risco por parte do tomador antes que lhe seja concedido o crédito. Essa sinalização pode ser obtida, por exemplo, pela exigência de uma garantia colateral. Para um tomador de alto risco, a garantia lhe trará um custo esperado alto, se comparado ao tomador de baixo risco, na medida em que a probabilidade de execução da garantia é maior. Esse sinal, segundo Malaga (1999), minimiza a inadimplência por parte do tomador.

Para selecionar tomadores e administrar os seus créditos, as empresas necessitam de, primeiramente, conhecer o risco do cliente. Tendo esta informação em mãos, outra alternativa, além da proposta por Malaga (1999), é a oferta de juros diferenciados para cada nível de risco. Essa solução, para as empresas que possuem modelos de risco desenvolvidos, adequadamente implementados e tecnologicamente suportados, permite reduzir seu nível de inadimplência, uma vez que é possível conceder e administrar melhor o crédito, além de gerenciar a carteira de inadimplentes.

Outra constatação que se pode fazer a respeito da inadimplência no Brasil é sua íntima ligação à atividade econômica, ou seja, se a economia do país está em condições satisfatórias, os índices de inadimplência são menores que num ambiente econômico desfavorável. Segundo Caouette; Altman; Narayanan (1999), quando a economia tem vigor, a onipresença do crédito pode parecer benigna. Em períodos de depressão econômica, contudo, pode se tornar uma preocupação mais séria. As pessoas perdem seus empregos e atrasam seus pagamentos. No Brasil, segundo pesquisa realizada em fevereiro de 2001 pelo Grupo Unidos (uma das maiores empresas especializadas em Cobrança e Crédito do país), o principal motivo que leva as pessoas a atrasarem os seus pagamentos são as compras para terceiros, que respondem por cerca de 21% da inadimplência, conforme ilustra a Figura 2.3.

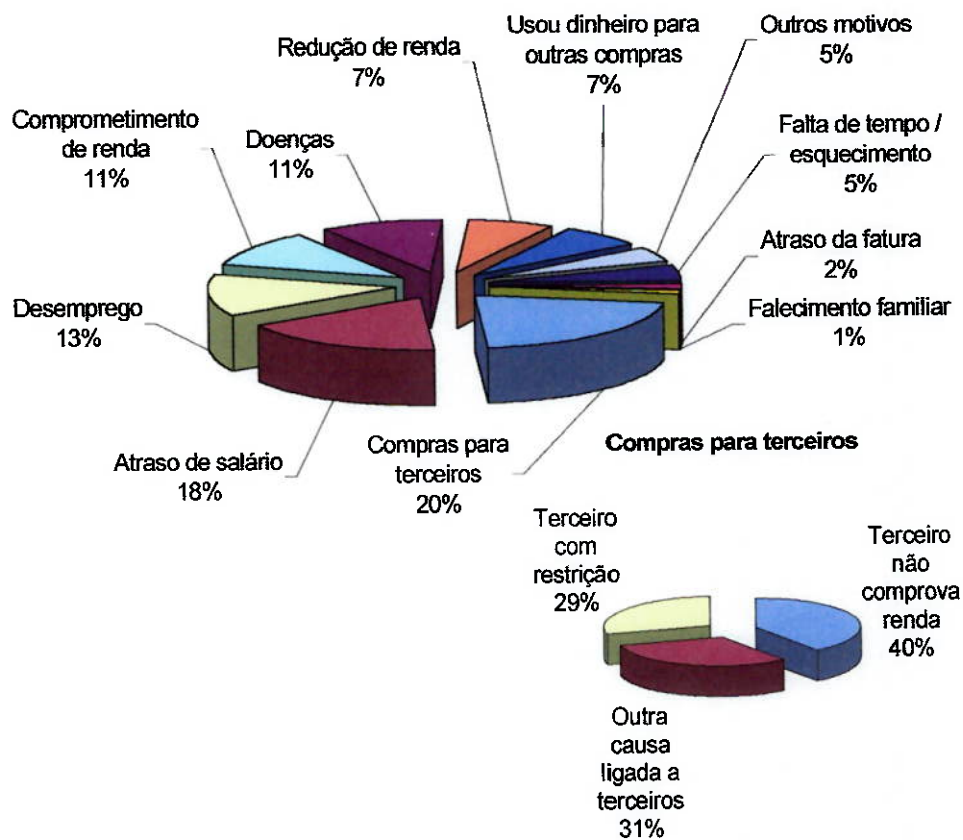


Figura 2.3 - Motivos do atraso nos pagamentos de prestações (em %).

Fonte: <<http://www.grupounidos.com.br>>

De acordo com a Tabela 2.4, para os devedores que estão em atraso há menos de 30 dias o principal motivo é o *Atraso de salário*, seguido de *Compras para terceiros* e *Falta de tempo / Esquecimento*. Já os que estão em atraso entre 31 e 90 dias, os três principais motivos são *Compras para terceiros*, *Atraso de salário* e *Comprometimento de renda*. Finalmente, para as dívidas com atraso superior a 91 dias o principal motivo é o *Desemprego*, seguido de *Compras para terceiros* e *Doenças*.

Até 30 dias		Entre 31 e 90 dias		Acima de 91 dias	
Atraso de Salário	25,8	Compras para terceiros	23,4	Desemprego	24,0
Compras para terceiros	20,3	Atraso de Salário	17,0	Compras para terceiros	19,9
Falta de tempo / Esquecimento	9,1	Comprometimento de renda	12,6	Doenças	13,0
Demais	44,8	Demais	47,0	Demais	43,1

Tabela 2.4 - Motivos da inadimplência por tempo do atraso(em %).

Fonte: <<http://www.grupounidos.com.br>>

Além destes dados obtidos sobre os motivos da inadimplência, outra pesquisa, esta realizada pela Associação das Empresas de Recuperação de Crédito (ASERC) em 2000, nos traz um importante resultado: de acordo com a Figura 2.4, 52% da procura por empréstimos refere-se ao pagamento de outras dívidas, uma vez que os consumidores sabem que se deixarem de pagar alguns serviços essenciais, como telefone ou eletricidade por exemplo, estes serão interrompidos de imediato.

Nesse sentido, o cartão de crédito é uma das principais “vítimas” da inadimplência, na medida em que o único empecilho para o consumidor, num primeiro momento, é o bloqueio do seu cartão, evitando que ele o utilize para mais compras. Não há, neste mercado, exigências de garantias ao se conceder o crédito e tampouco o corte de serviços essenciais que levem o consumidor com dificuldades financeiras a priorizar o pagamento do cartão.

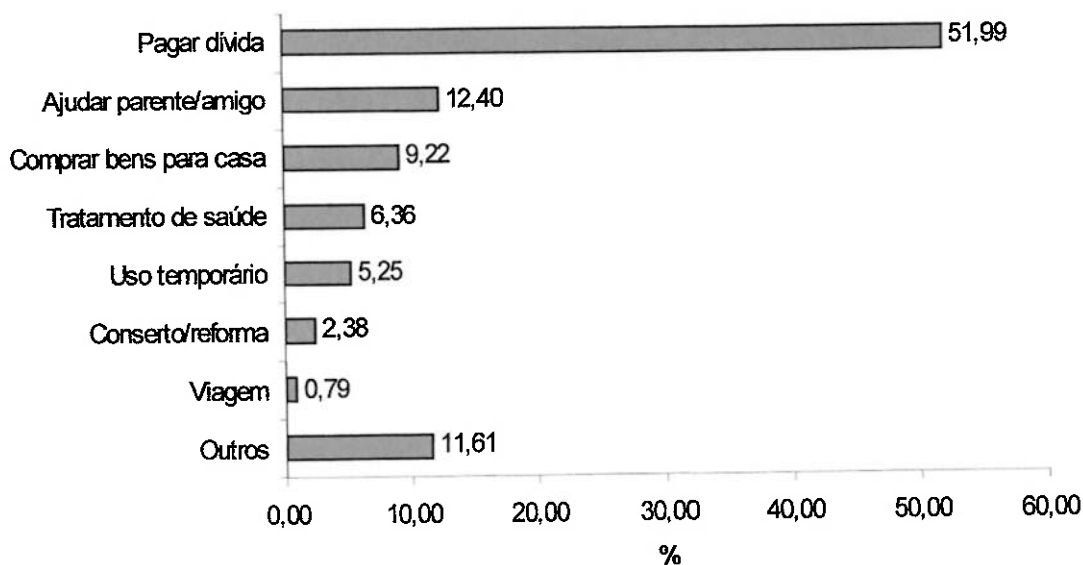


Figura 2.4 – Motivos da procura pelos empréstimos (em %). Fonte: Associação das Empresas de Recuperação de Crédito (ASERC). Pesquisa das carteiras em jan/2000.

Obs.: Não inclui veículos.

2.2.1 Inadimplência no Negócio de Cartões de Crédito

A indústria de cartões de crédito sentiu os efeitos da restrição ao crédito e do aumento da taxa de juros que ocorreram no segundo semestre de 1995 e primeiro semestre de 1996. Conforme mencionado anteriormente, as medidas restritivas tinham como objetivo adequar o ritmo de crescimento da demanda ao programa de estabilização do governo após a implantação do Plano Real, que havia proporcionado um forte crescimento do crédito para financiar o consumo.

Um dos efeitos das medidas restritivas foi o aumento da inadimplência. Empresas administradoras de cartões de crédito que conviviam com um índice de perda de crédito de cerca de 2%, viram este índice aumentar para aproximadamente 3,5% a partir do segundo semestre de 1996 até o final de 1998. Nesse período, as administradoras começaram a restringir ainda mais o crédito, através da adoção de critérios mais seletivos para a concessão de cartões.

São considerados como perda de crédito os valores não recuperados passados 180 dias da data da obrigação. Grande parte dessa perda é decorrente da falta de capacidade de pagamento do usuário de cartão de crédito, que optou por utilizar o crédito rotativo disponível (apresentado no Seção 2.1.4 deste trabalho). Estima-se que, do total de usuários de cartões de crédito no país, entre 40 e 45% tem propensão a financiar seus gastos (utilização do crédito rotativo). O índice de inadimplência, medido pela proporção de clientes que não conseguem pagar suas dívidas em até 10 dias após seu vencimento gira em torno de 10%, enquanto que a média de perda de crédito das administradoras de cartões de crédito no Brasil é de aproximadamente 2%.

Como já discutido, no mercado de cartões de crédito não é exigida garantia alguma pelo não pagamento da fatura mensal. Assim, o que cada administradora busca fazer para evitar a inadimplência é oferecer mais e melhores serviços aos consumidores, de modo a convencê-los de que a utilização do cartão lhes trará benefícios. Para que isto aconteça, é fundamental que o usuário do cartão mantenha seus pagamentos em dia, mesmo que seja utilizando seu limite de crédito rotativo. Caso contrário, seu cartão de crédito é bloqueado e os serviços essenciais cortados, assim que atinja 10 dias sem pagamento de sua fatura.

O resultado de uma pesquisa, realizada durante os meses de junho a agosto de 2001, por um dos escritórios de cobrança que presta serviços para a empresa aqui estudada, aponta que os principais motivos para o não pagamento das dívidas com cartões de crédito são, na ordem: o desemprego, o financiamento para terceiros, problemas de saúde na família, discordância do valor cobrado e problemas com a mercadoria financiada, de acordo com a Figura 2.5.

Estas principais causas diferem, com alguma significância, daquelas apresentadas anteriormente, que contemplavam todas as operações de crédito, sobretudo no que diz respeito à discordância com o valor cobrado, que aparece aqui com acentuada significância, devido, quase que exclusivamente, aos altos índices de fraude que são registrados neste mercado e aos juros do crédito rotativo. Seja através da compra com

um cartão “clonado”³, da chamada “auto-fraude”⁴, ou mesmo do juros cobrados, o associado não reconhece sua dívida e não efetua o pagamento. A questão dos problemas com a mercadoria financiada é outro fator específico deste negócio, pois o cliente não paga diretamente ao lojista, este o faz ao administrador do cartão com algum intervalo de tempo em relação ao momento da compra. Isto permite que o associado usufrua da mercadoria, que, caso apresente qualquer deformidade, desperta no cliente, o desejo de questionar ou mesmo não quitar a dívida.

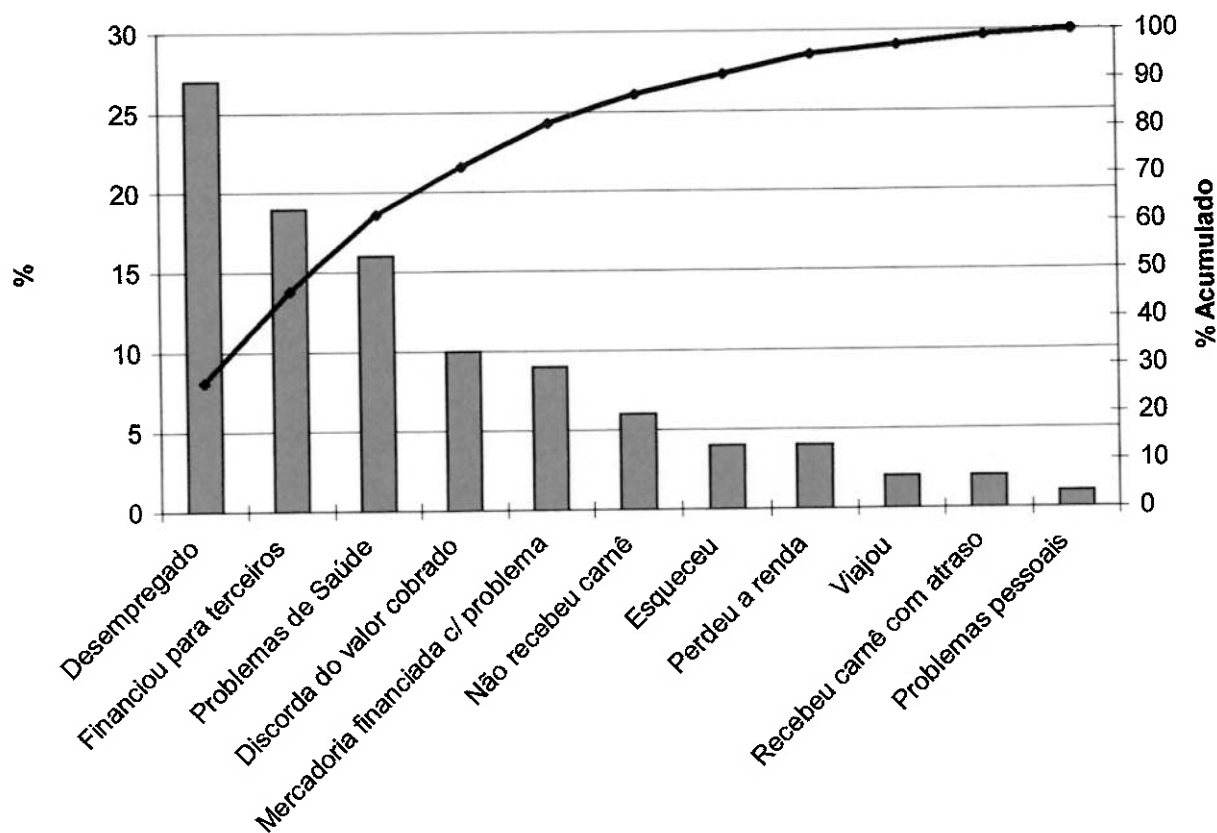


Figura 2.5 – Motivos de Inadimplência no Cartão de Crédito. Fonte: Tec Cobra – Escritório de Cobrança. Pesquisa de inadimplência de agosto de 2001.

³ A clonagem de cartões, mais conhecida como *skimming*, é uma modalidade de fraude extremamente comum entre as administradoras de cartões de crédito. Através de equipamentos cada vez mais sofisticados, o fraudador copia a tarja magnética do cartão, confecciona um novo e o utiliza para compras.

⁴ A “auto-fraude” consiste da compra de mercadorias por parte do associado que, livrando-se de seu cartão, entra em contato com a administradora afirmando que teve seu cartão perdido ou roubado e, conseqüentemente, aquelas compras não foram feitas por ele.

O fato do cartão de crédito não ser uma das principais prioridades de pagamento dos consumidores com dificuldades financeiras, torna-o mais suscetível à inadimplência frente a outras formas de aquisição de crédito, como por exemplo, o crédito imobiliário que, pela falta de pagamento, pode até mesmo não liberar ou retirar o imóvel de um consumidor com dívida em atraso.

De modo a se proteger contra esta suscetibilidade à inadimplência, as administradoras de cartões de crédito procuram ser extremamente criteriosas na concessão do crédito e elevar o *spread* (diferença entre a taxa de juro paga pelo banco para captar dinheiro e a taxa cobrada pelo seu empréstimo) para cobrir a despesa com devedores duvidosos (provisão regulamentada pela Resolução 2682 do Banco Central, descrita na Seção 2.3.3). Por essa razão, os juros do cartão de crédito são hoje os mais elevados do Brasil.

Assim, algumas administradoras têm percebido a importância de se trabalhar, de forma estratégica, a gestão do risco de crédito. Mais do que isso, uma vez tendo ciência de que a inadimplência é uma realidade, cabe às administradoras desenvolver políticas e estratégias para gerenciar da melhor forma sua carteira de inadimplentes e prejuízo fazendo com que a recuperação de perdas tenha relevância diante do resultado global da organização.

2.3 Cobrança

Cobrança, ou como preferem alguns, Recuperação de Crédito é o nome designado à área de uma companhia de cartões de crédito responsável por administrar a carteira de créditos vencidos, ou seja, de clientes em inadimplência.

É de responsabilidade da Cobrança, todo o processo de informação, contato e tentativa de conversão dos associados inadimplentes, novamente em associados “em dia”, preferencialmente antes que suas contas atinjam 90 dias em atraso, e conseqüentemente sejam canceladas. Este processo de evitar o cancelamento por inadimplência das contas é de fundamental importância para as administradoras, pois,

com a competição extremamente acirrada por clientes existente no mercado atualmente, a perda de clientes é significativamente sentida por qualquer dos competidores.

Este processo se inicia com a prevenção feita sobre os clientes com alta probabilidade de inadimplir, passando pelo envio de cartas, contatos telefônicos, envio dos dados do associado aos órgão de proteção ao crédito e chegando até o acionamento de escritórios externos de cobrança, cujo processo está detalhado no Seção 2.3.1 seguinte.

2.3.1 Formas de Cobrança

A administradora de cartões de crédito, na qual este estudo baseou-se, divide suas ações de cobrança em quatro maneiras distintas:

- cobrança através de cartas;
- cobrança através de contato telefônico;
- cobrança através de cobradores externos;
- cobrança de saldos altos.

Detalharemos, a seguir, cada uma delas de modo a explicitar de que se tratam e em quais momentos, dentro do período em que o cliente se encontra em inadimplência, elas atuam. Cumpre ressaltar que as maneiras citadas podem ser aplicadas em paralelo, dependendo da situação na qual se encontra o cliente inadimplente.

Cobrança através de cartas

De acordo com o conteúdo das correspondências que são enviadas aos associados que se encontram em inadimplência, podemos subdividir esta forma de cobrança em dois subgrupos: cobrança através de cartas informativas e cobrança através de cartas proposta.

A primeira consiste em uma correspondência enviada ao cliente assim que este atinge o décimo dia sem pagamento, pelo menos do valor mínimo de sua fatura, e tem

seu cartão bloqueado para utilização. Nesta carta, o cliente é informado sobre sua dívida vencida, ressaltando o valor devido, o período de atraso desde a data do vencimento e a maneira como este associado poderia regularizar sua situação, desbloqueando, conseqüentemente, seu cartão. O outro tipo de correspondência é aquele destinado aos associados que já se encontram com sua fatura atrasada há mais de 55 dias, período em que, de acordo com a política de cobrança da instituição em questão, não se permite financiar a dívida do associado. Esta carta, que é enviada em formato de ficha de compensação bancária para facilitar o primeiro pagamento, oferece ao associado 10 formas distintas de parcelamento de sua dívida.

Cobrança através de contato telefônico

Tendo um cliente não pago sua fatura em até dez dias decorridos de seu vencimento, uma central de operadores especialmente treinados em negociação é acionada para realizar a cobrança telefônica deste associado. Através de ligações contundentes, os negociadores buscam, de uma forma elegante e sem constrangimento, intimidar o cliente de modo que este prometa o pagamento total (ou mesmo parcial, nos casos em que ainda seja possível, de acordo com a quantidade de dias em atraso, efetuar o pagamento do valor mínimo da fatura) de sua dívida, ou negocie um acordo verbal de financiamento.

Esta central telefônica direciona seus esforços em duas atividades: a cobrança ativa e a cobrança receptiva. A cobrança ativa é aquela que fará o primeiro contato com o associado inadimplente, orientando-o e alertando-o sobre as conseqüências de sua permanência como inadimplente, ou seja, este seu caráter de iniciativa é o que a diferencia da receptiva, a qual apenas recebe os telefonemas dos associados que têm interesse em negociar sua dívida.

Cobrança através de cobradores externos

Assim que o atraso da fatura de um cliente ultrapassa os 180 dias, este tem sua conta encaminhada a uma empresa terceirizada (escritório externo especializado em cobrança) para que tenha continuidade a tentativa de recuperação do crédito.

Estes escritórios têm liberdade para adoção de mecanismos próprios para a operação de cobrança das contas em inadimplência, sendo somente restringidos pela política de recuperação de crédito da empresa proprietária da carteira. Essa política estabelece as condições para renegociação da dívida, os parcelamentos e descontos, de acordo com a quantidade de dias em atraso. Cabe aqui ressaltar que este tipo de cobrança exige grandes esforços de auditoria e fiscalização das empresas contratadas, de modo que os princípios éticos que regem a atividade de cobrança sejam respeitados, assim como o Código de Defesa do Consumidor.

Cobrança de Saldos Altos

Excetuando-se as dívidas caracterizadas como fraude, existe um grupo de clientes devedores que demanda uma cobrança diferenciada. Estes clientes cuja dívida supera os 30 mil reais são, em geral, clientes VIPs ou com renda muito acima da média da carteira de devedores.

Desta forma, a atividade de recuperação destes créditos passa a merecer uma atenção especial, sendo realizada por uma célula da área de cobrança, especializada na tentativa de recebimento de saldos devidos superiores ao valor mencionado acima, qualquer que seja o tempo em que o cliente se encontre em atraso.

Inclusão dos dados do associado nos órgãos de proteção ao crédito

Finalmente, em meio a todas estas formas distintas de cobrança mencionadas acima, a administradora de cartões de crédito dispõe de mais uma ferramenta extremamente poderosa e eficaz na tentativa de recuperar seus créditos vencidos: trata-se

da inclusão dos dados do associado nos órgãos de proteção ao crédito, a qual prejudica a pessoa física na obtenção de outros empréstimos e na efetivação de outras compras.

2.3.2 Provisão para Devedores Duvidosos

Nem todo o recebimento de valores devidos é líquido e certo, uma vez que as empresas estão sujeitas aos riscos de crédito. Caso a empresa tenha perdas com seus clientes (os não bons pagadores), seu saldo de Contas a Receber será reduzido, ou seja, a organização não receberá o montante registrado, mas este valor subtraído das possíveis perdas. Dessa forma, podemos identificar dois valores referentes às Contas a Receber: aquele registrado como contrapartida aos créditos concedidos e aquele que efetivamente a instituição irá receber (já deduzidas as possíveis perdas). Assim, em termos contábeis, a Provisão para Devedores Duvidosos (ou PDD) é constituída exatamente por esta diferença.

Dentro do Balanço Patrimonial das empresas, a PDD trata-se de uma conta redutora do Ativo, uma vez que a importância dedutível como provisão para créditos de liquidação duvidosa será a necessária a tornar a provisão suficiente para absorver as perdas que provavelmente ocorrerão no recebimento dos créditos existentes.

Sua constituição para instituições financeiras, é regulamentada pela Resolução 2682 do Banco Central, que está apresentada na Seção 2.3.3 a seguir.

2.3.3 Resolução 2682 do Conselho Monetário Nacional

Desde o mês de março do ano 2.000, passou a vigorar a Resolução 2.682, do Conselho Monetário Nacional, que estabeleceu novas regras para o provisionamento de créditos em liquidação, baseadas na avaliação da qualidade das carteiras dos bancos. Dessa forma, os bancos tiveram que classificar as operações de crédito utilizando critérios verificáveis, fundamentados pelas informações internas e externas sobre os seus clientes, tais como capacidade de pagamento, garantias, prazos e outras.

Dentre as principais disposições desta resolução, podemos destacar:

- Classificação das operações de crédito, em ordem crescente de risco, em um dos 9 níveis apresentados na Tabela 2.5 a seguir:

Nível de Risco	Atraso (dias)	PDD
A	até 14 dias	0,5%
B	15 a 30	1%
C	31 a 60	3%
D	61 a 90	10%
E	91 a 120	30%
F	121 a 150	50%
G	151 a 180	70%
H	acima de 180	100%

Tabela 2.5 – Regra de Aprovisionamento – Adaptada da Resolução 2682 do Banco Central do Brasil

- A classificação das operações de crédito de um mesmo cliente ou grupo econômico deve ser definida considerando aquela que apresentar maior risco e;
- Revisão periódica dos níveis de risco.

Esta resolução é encontrada, na íntegra, no Anexo I deste trabalho.

3 EMBASAMENTO TEÓRICO

Esta etapa do trabalho visa apresentar os fundamentos teóricos necessários para a construção e o entendimento de um modelo estatístico de predição. Serão explanadas, uma a uma, todas as técnicas estatísticas que compõem o conjunto de conhecimentos utilizados para a construção do modelo.

As Técnicas Estatísticas compreendem um conjunto de conceitos e ferramentas utilizadas no planejamento da coleta de dados, na sua análise e desenvolvimento de modelos de predição visando resolver problemas reais. Sua aplicação em situações ou processos em que a incerteza e a variabilidade estão presentes, torna estas técnicas ferramentas de extrema relevância e de grande utilidade.

3.1 *Estatísticas Descritivas*

Quando avaliamos um processo, não estamos interessados nos valores individuais de cada cliente, mas sim, no padrão de comportamento como um todo. O padrão pode ser caracterizado por alguns poucos números e gráficos que quantificam e exibem informações importantes.

A sumarização dos aspectos importantes de um conjunto de dados é chamada de Estatística Descritiva. Sob esse nome reunimos um conjunto de técnicas que visam condensar os dados através de tabelas, e calcular alguns números que indicam a tendência central dos dados (valor médio), e outros que medem a variabilidade dos mesmos (dispersão).

Desta forma, analisamos cada conjunto de dados (variáveis) sob os seguintes aspectos:

- a existência de observações atípicas que pareçam se destacar do conjunto dos dados e;
- as medidas numéricas que indiquem o valor central e quantifiquem a variabilidade presente nos dados.

Com isso, introduziu-se dois conceitos importantes no estudo das características numéricas, e que são intensamente utilizados: os conceitos de localização e de dispersão (ou variabilidade).

A localização pode ser caracterizada de várias formas:

- Mínimo: o menor valor observado;
- Máximo: o maior valor observado;
- Média aritmética: o valor que representa o centro da distribuição dos valores;
- Mediana: valor alternativo à média aritmética para representar o centro da distribuição, pois esta corresponde ao valor que cuja metade das medidas é menor ou igual a ele e a outra metade é maior ou igual a ele;
- Moda: valor que ocorre mais freqüentemente na distribuição.

A dispersão ou variabilidade é caracterizada por:

- Amplitude: a diferença entre o maior e o menor valor do conjunto de dados;
- Desvio Padrão: definido como $s = \sqrt{\frac{\sum_{x=i}^n (x_i - \bar{x})^2}{n-1}}$ sendo x_i os valores da variável, n o número de observações (ou tamanho da amostra) e \bar{x} a média das observações;
- Variância amostral: corresponde ao quadrado do desvio padrão.

3.2 *Análise de Correlação*

A análise de correlação mede o grau de associação entre duas variáveis independentes que poderão compor um modelo. Supondo haver duas variáveis preditoras X e Y, a correlação entre estas será positiva se valores altos de X estiverem associados a valores altos de Y. Caso ocorra o inverso, isto é, valores altos de X estejam associados a valores baixos de Y, a correlação será negativa. Estes valores de correlação podem variar entre -1 e 1 e, segundo Costa Neto (1977), a partir do fator |0,6| a correlação pode ser considerada alta e passível de análise.

Assim, o objetivo desta análise é estudar inter-relações entre as variáveis, identificando através da matriz de correlação construída, variáveis com forte correlação, o que corresponderá a sobreposição de informações e, neste caso, pode-se diminuir o número de variáveis a serem analisadas na elaboração do modelo.

3.3 Definição da Técnica Estatística

A modelagem dos dados aqui em questão visa identificar matematicamente uma ligação entre uma variável resposta (Y) e as variáveis chamadas preditoras (X_n) da forma como está exposta a função a seguir:

$Y = a_1X_1 + a_2X_2 + \dots + a_kX_k$, onde X_1, X_2, \dots, X_k são as variáveis preditoras (ou independentes) e a_1, a_2, \dots, a_k são seus pesos para a definição da variável Y resposta.

Todas estas variáveis envolvidas, sejam elas independentes (entrada) ou de resposta (saída), podem ser qualitativas ou quantitativas (discretas ou contínuas respectivamente). “As quantitativas são atributos que podem ser associados a uma escala de valores e conseqüentemente comparados matematicamente (maior, menor, etc.) e as qualitativas representam atributos que não podem ser medidos numericamente, como por exemplo raça, sexo, religião, etc.” (Nakamura, 2000, p. 40 e 41).

A aplicabilidade das técnicas existentes é o fator principal que as diferencia, ou seja, os tipos de variáveis de entrada e saída aos quais é indicada determinada técnica permite identificar as diferenças entre elas. Assim sendo, a Tabela 3.1 apresentada a seguir foi utilizada como base para a escolha da técnica aplicada:

Técnica Estatística	Variável independente		Variável resposta	
	Qualitativa	Quantitativa	Qualitativa	Quantitativa
Correlação Canônica	X	X	X	X
Análise de Variância Multivariada	X			X
Análise de Variância	X			X
Análise de Discriminante Múltiplo		X	X	
Regressão Logística	X	X	X	
Análise de Regressão Múltipla	X	X		X
Análise Conjunta	X		X	X
Modelagem Estrutural de Equações	X	X		X
Análise de Agrupamento		X	-	-
Análise Fatorial		X	-	-

Tabela 3.1 – Relação entre as variáveis (Análise multivariada) – Adaptada de Hair (1995) apud Nakamura (2000)

Conforme pode-se perceber da Tabela 3.1, a regressão logística permite variáveis independentes tanto qualitativas quanto quantitativas, como é o caso das variáveis que serão utilizadas na confecção deste modelo (saldo devedor, tempo de relacionamento do cliente com a instituição, tipo de produto, bandeira, etc.). Paralelamente a isto, nota-se através da mesma tabela que esta técnica é especificamente adequada para casos onde a variável resposta é qualitativa (conta com saldo devedor contabilizado como prejuízo ou não), além de possuir a principal vantagem de não ser muito afetada pelo comportamento do conjunto de dados (principalmente para amostras cuja distribuição não é normal).

Diante de tudo o que foi mencionado, decidiu-se aplicar a regressão logística na modelagem dos dados para a construção do modelo de predição de contas que não serão contabilizadas como prejuízo (com mais de 180 dias de atraso). A Seção 3.4 exibe maiores informações sobre a técnica escolhida.

3.4 Regressão Logística

Ao tratarmos de problemas multivariados, deparamo-nos com a situação em que se deseja modelar a probabilidade de um evento em função de um conjunto de variáveis independentes. Estas variáveis, que aqui chamaremos de X , podem ser contínuas ou categóricas e, além disso, os eventos são dicotômicos, isto é, a variável resposta, que denominaremos Y , deve ser SIM ou NÃO. Resumindo, $Y = f(X_1, X_2, \dots, X_n)$ onde $f(X_1, X_2, \dots, X_n)$ é uma função linearizável.

Usualmente, ao evento SIM é atribuído o número 1 e ao evento NÃO é atribuído 0 (zero). Dessa forma, passamos a trabalhar com uma resposta Y que pode assumir os valores 0 ou 1.

Sendo $\pi = \text{Prob}(Y = 1)$, isto é, a probabilidade de um indivíduo escolhido aleatoriamente assumir o valor 1 para a variável resposta, teríamos o seguinte modelo de regressão, caso utilizássemos as técnicas de regressão linear:

$\pi = \pi(Y = 1) = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k$, onde X_1, X_2, \dots, X_k são as variáveis independentes.

Contudo, este tipo de modelo não pode ser utilizado, já que não podemos garantir que os valores preditos por ele estejam entre 0 e 1. Para garantirmos a satisfação desta condição é que adotamos o Modelo Logístico, pois sua técnica de ajuste e análise, a Regressão Logística, assegura que os valores preditos de π estarão sempre entre 0 e 1, conforme podemos perceber a seguir.

3.4.1 Análise Bivariada: Probabilidade e Odds

Consideremos o seguinte exemplo: uma empresa de cartão de crédito deseja determinar se uma proposta para aquisição de um cartão de crédito é fraudulenta ou não, baseada no fato da data de nascimento informada pelo proponente estar diferente da apresentada no cadastro das empresas telefônicas.

Adotamos o valor 1 para indicar propostas fraudulentas (F), o valor 0 para propostas não fraudulentas (N), valor 1 no caso da data ser diferente (D) na base das empresas telefônicas e 0 caso seja igual (I). Uma amostra de 24 propostas resultou na Tabela 3.2:

Propostas Fraudulentas	Data de nascimento	Propostas Não Fraudulentas	Data de nascimento
1	1	0	1
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	1	0	0
1	0	0	0
1	0	0	0

Tabela 3.2 – Propostas e comparação da data de nascimento

Os dados da Tabela 3.2 foram sintetizados e são, desta forma, apresentados na Tabela 3.3:

Proposta	Data de nascimento		Total
	Diferente (D)	Igual (I)	
Fraudulenta (F)	10	2	12
Não fraudulenta (N)	1	11	12
Total	11	13	24

Tabela 3.3 – Tabela de Contingência

Da Tabela 3.3 temos, então, que:

- a probabilidade de ser fraude é 0,50, pois temos 12 fraudes em 24 propostas;
- a probabilidade de ser fraude dado que a data de nascimento é diferente é 0,909, pois temos 10 fraudes em 11 propostas com data diferente e;
- a probabilidade de ser fraude dado que a data de nascimento é igual é 0,154, pois temos 2 fraudes em 13 propostas com data igual.

Isto posto, usualmente, em regressão logística usamos uma medida denominada razão de *Odds*, que torna relativas as várias probabilidades de interesse, ou seja, a razão de *Odds* corresponde à relação entre a probabilidade π de um evento ocorrer pela probabilidade $1-\pi$ de não ocorrência deste mesmo evento $\left(\frac{\pi}{1-\pi}\right)$.

No exemplo, temos, respectivamente, que:

- a razão de *Odds* da proposta fraudulenta é 1, pois temos 12/12;
- a razão de *Odds* da proposta fraudulenta dado que a data de nascimento é diferente é 10, pois temos 10/1 e;
- a razão de *Odds* da proposta fraudulenta dado que a data de nascimento é igual é 0,182, pois temos 2/11.

3.4.2 Modelo de Regressão Logística

O modelo logístico, de modo semelhante à regressão linear múltipla, visa determinar a variável resposta Y binária, a partir das variáveis independentes X , ou seja:

$$Y = f(X_1, X_2, X_3, \dots, X_k), \text{ onde } f(X_1, X_2, X_3, \dots, X_k) \text{ é uma função linearizável.}$$

Sabendo-se que π é a probabilidade de $Y = 1$, um modelo bastante usado para estudar o comportamento de Y em função de $X_1, X_2, X_3, \dots, X_k$ é o modelo de regressão logística, dado pela função $g(x)$ (chamada *logit*), que é combinação linear das variáveis explicativas adicionadas de uma constante, conforme segue:

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k \text{ onde, no exemplo dado } \pi \text{ é a}$$

probabilidade de uma proposta ser fraudulenta.

Através de uma transformação algébrica obtemos a seguinte equação para $\pi(x)$:

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

Assim, $\pi(x) = \frac{1}{1 + e^{-\ln(\text{odds})}}$ e sendo $f(x) = \frac{1}{1 + e^{-\ln(\text{odds})}}$ temos a seguinte curva

(Figura 3.1):

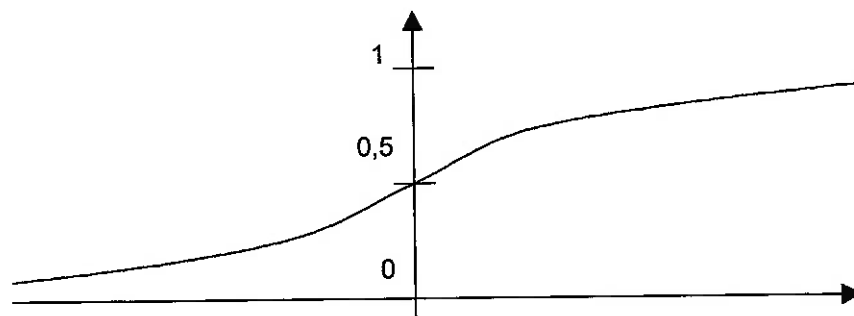


Figura 3.1 – Gráfico de $f(x) = \frac{1}{1 + e^{-\ln(\text{odds})}}$

Conforme mencionado anteriormente, o modelo de Regressão Logística é utilizado em situações onde os valores preditos devem estar entre 0 e 1. Com esta função cujo Conjunto Imagem está compreendido entre este intervalo, torna-se possível tal solução.

3.5 Adequação do Modelo

A modelagem tem como objetivo a determinação do peso de cada variável independente na composição da variável resposta, ou seja, determinar os valores do vetor $A = (a_0, a_1, a_2, \dots, a_k)$.

Como vimos, a relação entre a probabilidade π e as variáveis independentes do modelo não é linear, enquanto a relação entre o log do Odds e estas variáveis é linear. Com isso, a interpretação dos coeficientes das variáveis independentes é dada a partir do log do Odds e não da probabilidade π . Usualmente, os coeficientes são estimados com base no método da Máxima Verossimilhança. Esse método estima os coeficientes através da função chamada “Função da Verossimilhança”, onde os estimadores de máxima verossimilhança para esses coeficientes são determinados de maneira que esta função seja maximizada. Analiticamente, consiste em determinar a solução do seguinte sistema de equações:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0$$

$$\sum_{i=1}^n x_j [y_i - \pi(x_i)] = 0$$

Para se obter a solução deste sistema de equações, usualmente recorremos a métodos iterativos, os quais, como qualquer cálculo estimado, estão sujeitos a erros. Esta estimativa do erro padrão dos coeficientes obtém-se construindo-se a matriz de informação com os seguintes valores:

$$\frac{\partial^2 L(a)}{\partial a_j^2} = -\sum_{i=1}^n x_j^2 \pi_i (1 - \pi_i) \quad \text{e} \quad \frac{\partial^2 L(a)}{\partial a_j \partial a_u} = -\sum_{i=1}^n x_j x_u \pi_i (1 - \pi_i)$$

$$j, u = 0, 1, 2, \dots, p \quad \text{e} \quad \pi_i = \pi(x_i)$$

As variâncias e covariâncias do erro padrão das estimativas são dadas pelos elementos da matriz inversa de informação, sendo representadas respectivamente pelos elementos da diagonal e pelos elementos fora das diagonais.

Utilizaremos o SAS (*Statistical Analysis System*) para a solução dessas equações devido a sua complexidade. Ele fornece, além dos coeficientes e dos erros, o valor do logaritmo da Função Verossimilhança, o qual será utilizado nos testes de significância do modelo.

Dispondo dos coeficientes da equação, realizamos o teste de significância dos mesmos, ou seja, definimos se a variável relativa ao coeficiente realmente contribui para o resultado do modelo. Estes testes buscam reduzir o número de variáveis do modelo e as possíveis interações entre as mesmas, por consequência.

Para a seleção do melhor modelo, faremos uso do processo *Stepwise* de acordo com a metodologia fornecida por Hosmer e Lemeshow (1989). A importância das variáveis, de acordo com o método, é definida em termos de uma medida da significância estatística do coeficiente da mesma. Em se tratando da Regressão Logística, assumimos que os erros seguem uma distribuição de Bernoulli e sua significância é determinada através do teste Razão de Verossimilhança. Desta forma, em qualquer passo do *Stepwise*, a variável de maior importância será aquela que produzir a maior variação no valor da Função Verossimilhança se comparada a um modelo sem a variável.

O algoritmo do procedimento *Stepwise* é o seguinte:

- *Passo 0:* Define-se o valor da constante do modelo (caso em que todas as contas analisadas teriam a mesma probabilidade π de não serem contabilizadas como prejuízo) e calcula-se o valor de seu logaritmo de verossimilhança.
- *Passo 1:* Dentre todas as variáveis independentes consideradas para o estudo, escolhe-se aquela que melhor explica a variável resposta, ou seja, a variável independente x_1 que possui menor probabilidade p de seu coeficiente ser igual a zero. A partir disso, calcula-se o novo valor do logaritmo de verossimilhança L_i do modelo, com a inclusão da variável. Em seguida, obtém-se o valor da estatística G que representa, através de uma distribuição Qui-quadrado (X^2), a probabilidade do

coeficiente ser igual a zero e corresponde a $-2 (L_{i-1} - L_i)$, onde L_i trata-se do logaritmo de Verossimilhança da etapa i e L_{i-1} do logaritmo de Verossimilhança da etapa $i-1$. O valor de G encontrado é comparado com o valor crítico adotado com base no nível de significância definido: se G for maior que o valor crítico, introduz-se a variável no modelo e, em caso contrário, a variável não é incluída e encerra-se o algoritmo.

- *Passo 2:* Adiciona-se uma nova variável ao modelo e analisa-se a variação no valor do logaritmo de verossimilhança, repetindo-se o passo 1. Caso a variável tenha que ser incluída, testa-se a permanência da variável anterior, calculando-se a estatística G da seguinte forma: $G = -2 (L_{1x2} - L_2)$, onde L_{1x2} corresponde ao valor do logaritmo de verossimilhança do passo 1, considerando a inclusão da variável x_2 ao invés da variável x_1 e L_2 é o logaritmo de verossimilhança do passo 2 (com x_1 e x_2 no modelo). Caso este novo G seja maior que o valor crítico, a variável x_1 permanece, caso contrário ela é excluída.

O algoritmo se repete por mais vezes até que o critério de parada seja satisfeito, ou seja, até que o nível de significância adotado seja ultrapassado (para este trabalho, definiu-se nível de significância de 0,05).

3.6 Variáveis Dummy

Dentre as diversas variáveis independentes que podem colaborar na predição de um fenômeno, muitas vezes deparamo-nos com variáveis qualitativas (não numéricas) que, para poderem ser incluídas, necessitam de uma prévia codificação. Como exemplo, no mercado de cartões de crédito, podemos ter a bandeira do produto (Seção 2.1.2 deste trabalho), a qual assume 3 resultados possíveis: Visa, Mastercard ou outra. Esta variável qualitativa deve ser transformada em 2 variáveis dummy, de acordo com a Tabela 3.4:

Variável Principal Bandeira	Variáveis Auxiliares	
	X1	X2
Visa	0	0
Mastercard	0	1
Outras	1	0

Tabela 3.4 – Exemplo de codificação de variáveis qualitativas

Esta conversão permite identificar matematicamente a contribuição das variáveis qualitativas pois, por exemplo, se $X1 = 0$ e $X2 = 1$, temos para o produto, a bandeira Mastercard. Nota-se que para determinada variável independente que pode assumir k possibilidades, serão necessárias $k - 1$ variáveis auxiliares fictícias.

4 APLICAÇÃO DAS TÉCNICAS ESTATÍSTICAS

A partir desta etapa do trabalho, inicia-se a construção do modelo propriamente dito. Construir um modelo, significa definir uma equação matemática que contenha as variáveis que irão predizer a probabilidade de um cliente regularizar seu saldo devedor, antes que atinja 180 dias em atraso e tenha sua dívida contabilizada como prejuízo por parte da instituição; para tanto, serão atribuídos pesos a estas variáveis através da utilização de técnicas estatísticas descritas no Capítulo 3.

Embora haja técnicas estatísticas que definem quais variáveis devem fazer parte no modelo, esta inclusão das variáveis não se trata de uma decisão puramente quantitativa. McCahill (1998) afirma que todas as variáveis do modelo devem ser estatisticamente significantes, porém nem todas estas devem estar no modelo. Para aquelas cujo gerenciamento é extremamente caro ou que o tempo para obtenção das informações é alto, recomenda-se não utilizá-las no modelo, a menos que sua ausência diminua sensivelmente o poder de predição.

Portanto, a inclusão de variáveis no modelo de *charge off scoring* ocorreu através de uma combinação de discussões subjetivas com utilização de técnicas estatísticas. É importante que as pessoas envolvidas sintam-se confortáveis com as variáveis do modelo no que se refere a: aspectos legais, coerência, interpretação, sensibilidade a fatores econômicos e facilidade de gerenciamento.

4.1 Considerações Iniciais e Necessidade

Enquanto o índice de perda de crédito médio do mercado de cartões de crédito é da ordem de 2%, na instituição aqui estudada ele é bem menor¹. Uma das principais razões para isto foi o gerenciamento da carteira de inadimplentes como uma atividade estratégica para o resultado da empresa. Isso significou reassumir a administração de boa parte dos saldos das contas a receber em atraso, antes gerenciadas por empresas

¹ A confidencialidade da informação impediu que este índice fosse aqui divulgado.

terceirizadas (escritórios externos de cobrança). Enquanto no passado, as contas eram transferidas para as várias empresas terceirizadas imediatamente após atingirem 60 dias em atraso, hoje, estas empresas, em número bem menor, só recebem contas atrasadas há mais de 180 dias, ou seja, já contabilizadas como prejuízo.

Paralelamente a isto, mecanismos alternativos de cobrança e tecnologias adequadas foram desenvolvidas e implementadas, durante os últimos 24 meses, de forma rápida e eficaz, permitindo que os resultados internos de perdas de crédito estivessem absolutamente dentro dos padrões aceitáveis.

Partindo-se do pressuposto de que o cartão de crédito não está entre as principais prioridades de pagamento do consumidor com dificuldades financeiras, tal empresa foi inovadora ao adotar estratégias mais flexíveis de renegociação de dívidas, também chamadas de acordos. Essa estratégia torna-se vantajosa para ambas as partes: o devedor, muitas vezes, consegue maiores prazos para pagar, menores encargos e descontos em algumas situações, enquanto que o credor tem de volta o capital, seu maior bem, para investir em outros negócios.

Conforme apresentado anteriormente, a despesa financeira com provisão que cada conta causa à instituição está diretamente relacionada ao risco que esta oferece, classificado pelo Banco Central, o qual, por sua vez, se relaciona diretamente à quantidade de dias que o cliente permaneceu em atraso após o vencimento de sua fatura mensal. Assim, um cliente com 100 dias de atraso de sua fatura, devendo R\$1.000,00 representará uma despesa financeira com provisão equivalente a R\$300,00 (nível de risco E do Banco Central – entre 91 e 120 dias em atraso).

No momento em que este cliente efetua o pagamento total de sua dívida, a empresa deixa de contabilizar a parcela do saldo provisionada, eliminando suas despesas financeiras para esta conta. Algo semelhante ocorre quando o cliente paga a primeira parcela de um acordo para renegociação de sua dívida, pois assim, este deixa de estar em atraso com suas obrigações, evitando a aumento das despesas financeiras referentes à sua provisão para créditos de liquidação duvidosa, uma vez que seus dias de atraso são congelados até que o pagamento de alguma de suas parcelas do acordo atrase por um período de tempo superior, momento em que o risco da conta volta a aumentar e,

por conseqüência, o mesmo acontece com sua provisão, de acordo com as classificações determinadas pelo Banco Central na já citada Resolução 2682 (Anexo I).

Analisando-se exclusivamente os clientes com atraso superior a 91 e inferior a 120 dias (contas consideradas no desenvolvimento deste modelo estatístico), duas serão as ações tidas como produtivas de modo que se impeça o crescimento das despesas financeiras com provisão ou mesmo a contabilização de seus saldos como prejuízo: a quitação total da dívida do associado ou a renegociação efetiva de seu saldo devedor (entende-se por renegociação efetiva o pagamento da primeira parcela do acordo firmado entre as partes). Assim, buscando reduzir despesas com provisão e minimizar os custos internos, a empresa decidiu investir na construção um modelo estatístico que aponte os clientes com maior propensão a agir de acordo com quaisquer das duas formas supra citadas, clientes estes que não terão suas dívidas contabilizadas como prejuízo. A este modelo demos o nome de *Charge Off Scoring*, o qual é apresentado detalhadamente a partir da Seção 4.2.

4.2 O Modelo de Charge Off Scoring

O modelo de *Charge Off Scoring* é um modelo estatístico que tem como premissa o aprendizado com o comportamento passado de um cliente (ou melhor, de uma conta) em inadimplência, a fim de predizer a probabilidade deste cliente (ou conta) não ter seu saldo contabilizado como prejuízo. Assim, admite-se que as mesmas condições que separam contas “boas” e “más” no passado, se repetirão no futuro e que, por esta razão, pode-se predefinir que serão os novos bons e maus pagadores a partir da relação entre variáveis durante o relacionamento entre a instituição e o cliente.

A população a que este modelo se aplica é a de inadimplentes, ou seja, dado que este cliente atrasou mais do que 90 dias no pagamento de sua obrigações, este modelo indicará a probabilidade deste pagar totalmente ou negociar suas dívidas dentro de até 3 meses.

No modelo é suposto que as relações existentes entre as variáveis, ao longo do período de relacionamento do cliente com a instituição que lhe concedeu crédito, são diferentes para contas “boas” e “más”.

No cenário aqui considerado, define-se como conta “boa” aquela que é retirada da situação de inadimplência, negociando o financiamento ou pagando de uma só vez seu saldo devedor, e conta “má” aquela que não tem restabelecida sua situação original (cliente/conta em dia).

O principal benefício proveniente da utilização do modelo de *charge off scoring* é a otimização dos recursos de cobrança, reduzindo custos e consequentemente maximizando os resultados financeiros. Com esse modelo, prioriza-se as contas a serem cobradas, reduzindo o número de telefonemas desnecessários por parte da central de cobrança telefônica (no caso das contas “más”, que serão encaminhadas diretamente aos escritórios externos de cobrança, especialistas na recuperação de créditos vencidos com idade superior a 180 dias), diminui o atrito com clientes bons pagadores e aumenta a relação entre custo e benefício da infra-estrutura utilizada para gerenciar as contas a receber.

O objetivo do modelo de *charge off scoring* é fornecer, através de uma pontuação, uma medida que permita ordenar os clientes numa escala de menor para maior probabilidade de saldarem suas dívidas, antes que estas sejam contabilizadas como prejuízo.

Existem várias formas de se trabalhar com este tipo de modelo. O mais comum é considerar toda a população de inadimplentes e construir um modelo que diferencie o bom do mau, isto é, não procurar criar modelos distintos por faixa de atraso, produto, etc.

Entretanto, a construção de modelos para cada faixa de dias em atraso tem sido uma arma poderosa no aumento da eficiência dos modelos de cobrança, na medida em que aquela identifica as populações que apresentam relações de predição homogêneas. “Embora o processo de segmentação tenha sido criado baseado nas necessidades dos processos ou dos negócios, e não para aumentar a eficiência dos modelos de predição,

ele tem sido de muita utilidade” (Makuch et alia, 1998, p. 71)². Inicialmente, agrupa-se as populações de maneira que seja possível estabelecer um modelo para cada grupo de clientes em atraso e, a partir daí, o modelo de cada população é que, então, fará a separação entre as contas “boas” e “más”.

Os segmentos utilizados para o gerenciamento da inadimplência, pela instituição na qual realizou-se o estudo, são as faixas de atraso: de 1 a 30 dias, de 31 a 60 dias, de 61 a 90 dias e de 91 a 120 dias e acima de 120 dias. Para cada uma dessas faixas, constrói-se um modelo que irá prever a situação de cada cliente em um determinado período.

Como exemplo desta aplicação, pode-se citar a utilização dos modelos de cobrança para cada faixa de atraso do crédito imobiliário nos Estados Unidos. Cordell et al (1998) mostra a segmentação que é feita no gerenciamento da inadimplência nesse tipo de empréstimo: para o primeiro mês de inadimplência (1 a 30 dias), o modelo é chamado de *Early Indicator Collections Score* e, para as faixas além do primeiro mês de inadimplência (31 a 60 dias, 61 a 90 dias e 91 a 120 dias), os modelos são chamados de *Loss Mitigation Score*. O primeiro tipo de modelagem (*Early Indicator*) dá como resultado a probabilidade do empréstimo ultrapassar os 30 dias de atraso, considerados prioritários na estratégia de cobrança. Os modelos de *Loss Mitigation* fornecem a probabilidade dos empréstimos tornarem-se perda de crédito (prejuízo), ou seja, ordenam os empréstimos do mais baixo risco de perda para o mais alto, permitindo a elaboração de estratégias de recuperação para cada nível de risco.

4.2.1 Definições e Conceitos

Todos os modelos de risco de crédito, sejam eles de concessão ou de recuperação de crédito, têm como objetivo separar os “bons” dos “maus” clientes, normalmente traduzidos como bons e maus pagadores. Entretanto, para cada tipo de modelo de risco, a definição do que é bom e mau pode ser diferente: para os modelos de concessão de

² Tradução livre do autor.

crédito, por exemplo, o bom cliente é aquele que apresenta baixo risco de inadimplir, enquanto que para os modelos de recuperação, bons clientes são aqueles que apresentam alta probabilidade de pagamento, uma vez estando em inadimplência.

Contudo, este conceito de bom pode diferir, ainda que sutilmente, para os diversos modelos de recuperação de crédito. Exemplificando: se o cliente estiver com menos de 30 dias de atraso e efetuar o pagamento de sua dívida, este pode ser considerado bom; no entanto, se o único meio dele regularizar sua situação dentro desta faixa de atraso é fazer um parcelamento de seu saldo devedor, este será classificado como mau pagador, pois, de acordo com a política de recuperação de crédito desta empresa, só se justifica a negociação de um acordo com o associado, assim que este atinge 56 dias de atraso (isto se deve à diferença entre os juros cobrados pelo crédito rotativo, aproximadamente 12,90% ao mês, e os juros do financiamento da dívida que se mantém em 2,95% ao mês).

Por outro lado, caso o cliente se encontre na faixa de atraso de 56 a 90 dias e negocie um acordo, ele será classificado como bom, ao passo que aquele que nenhum pagamento realizou será mau.

Desta forma, podemos perceber o quanto o conceito de bom e mau pode diferir, de acordo com o objetivo que se deseja alcançar com a elaboração do modelo. Esse conceito pode justificar a necessidade de segmentação da carteira de inadimplentes antes da construção dos modelos de recuperação, uma vez que para cada classificação de risco (ou faixa de atraso) podem existir diferentes conceitos de bons e maus pagadores.

Diante disto, convencionou-se, para o modelo em questão, que os grupos de contas “boas” e “más” serão compostos da seguinte forma:

- Grupo Bom: Contas inadimplentes com atraso da fatura entre 91 e 120 dias que, após três meses, não atingiram 180 dias em atraso por terem saldado ou negociado o parcelamento de suas dívidas, ou seja, não tiveram seu saldo contabilizado como prejuízo.
- Grupo Mau: Contas inadimplentes com atraso da fatura entre 91 e 120 dias que, após três meses, atingiram 180 dias em atraso por não terem saldado ou negociado o

parcelamento de suas dívidas, ou seja, tiveram seu saldo contabilizado como prejuízo.

4.2.2 Desenho conceitual do modelo

Foram selecionadas, para o desenvolvimento do modelo, as contas com atraso entre 91 e 120 dias em T, observando-se, adicionalmente aos dados deste mês T, o histórico do mês T-1 e a situação de atraso da conta inadimplente em T+3. A Figura 4.1 ilustra claramente este desenho conceitual:

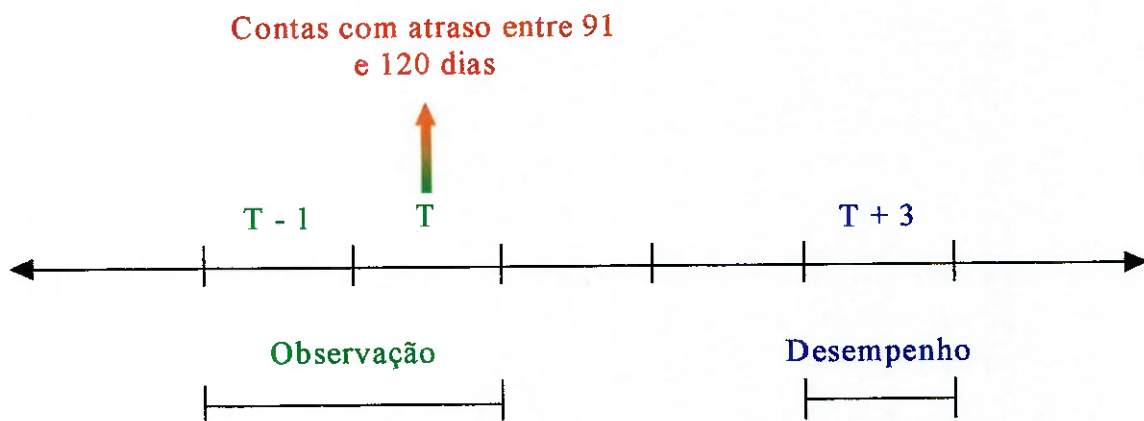


Figura 4.1 – Desenho Conceitual do Modelo

4.2.3 Delineamento Amostral

A coleta de dados para a composição de uma amostra, que represente significativamente a população estudada e para a qual deseja-se que o modelo se aplique, é o ponto de partida para todos os modelos estatísticos. Por esta razão, o delineamento amostral é um passo chave na construção deste modelo de *charge off scoring*, pois o sucesso ou fracasso da implementação pode ser definido pela qualidade dos dados usados na elaboração do modelo.

Assim, para a composição das amostras tanto de desenvolvimento quanto de validação do modelo, foram utilizadas informações provenientes de cinco cortes (meses). Os meses escolhidos foram os de setembro, outubro, novembro, dezembro de 2000 e janeiro de 2001. Conforme mencionado na Seção 4.2.2, adicionou-se às informações de cada corte, o comportamento da conta no mês anterior (T-1), conforme ilustrado na Figura 4.1. A Figura 4.2, a seguir, descreve claramente aquilo que foi executado:

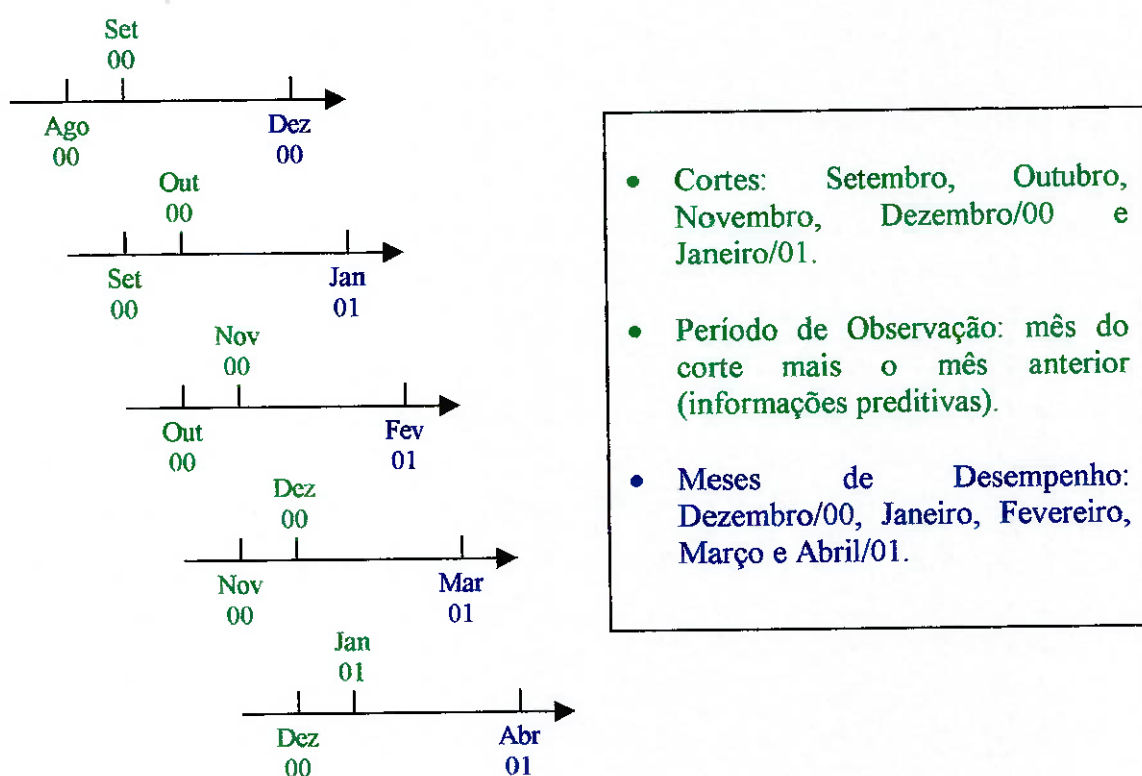


Figura 4.2 – Representação da Composição da Amostra

A unidade amostral deste estudo é a conta de cada cliente, nada impedindo que um mesmo cliente possua mais de uma conta dentro daquelas que foram amostradas. Com isso, o resultado da probabilidade de pagamento é para a conta, e não para o cliente.

Para o período considerado (setembro de 2000 a janeiro de 2001), foram selecionadas aleatoriamente 17.013 contas que atingiram os 91 dias sem a quitação de seus débitos. Destas contas, verificou-se que 68,71% delas continuaram em atraso 3 meses depois (“más”) e 31,29% delas negociaram ou pagaram totalmente suas dívidas (“boas”). Além disso, mais 3.249 contas foram selecionadas para compor a base de validação do modelo, sendo 2.191 “ruins” e 1.058 “boas”. A Tabela 4.1 resume o que foi dito.

Bases de Dados Amostrais	Contas “Más”		Contas “Boas”		Total	
	N	%	N	%	N	%
Base de desenvolvimento	11.689	68,71%	5.324	31,29%	16.843	100%
Base de validação	2.191	67,44%	1.000	32,56%	3.249	100%
Total	13.880	68,50%	6.382	31,50%	20.262	100%

Tabela 4.1 – Composição das bases de dados amostrais

De acordo com a política de crédito e cobrança da instituição na qual foi desenvolvido o modelo, as contas são canceladas ao atingir 91 dias em atraso do pagamento da fatura mensal, por se tratarem de clientes que não despertam mais o interesse de manutenção de relacionamento por parte da empresa; assim, todas as contas envolvidas no estudo são contas já canceladas pela instituição.

Cumprе ressaltar, também, que foram selecionadas para o modelo somente contas com utilização normal, ou seja, contas em atraso cujo saldo devedor é composto efetivamente por compras e/ou saques realizados com o cartão, excluindo-se aqui aquelas cuja dívida é representada apenas por taxa de anuidade e/ou serviços. Isto se dá por causa da política de recuperação de crédito da empresa estudada impedir a cobrança de dívidas destas naturezas.

As contas em inadimplência com saldo inferior a R\$100,00 também foram excluídas do estudo, pois, assim como as contas mencionadas logo acima, se encaixam

no grupo daquelas que não devem ser cobradas, de acordo com a política desta organização.

Além das exclusões citadas acima, as contas cujos associados negociaram acordo com a cobrança não tomaram parte neste estudo, por se tratarem de contas que se encontram em uma situação diferente daquelas aqui abordadas. Assim que um associado, que possui uma conta em atraso, paga a primeira parcela de um acordo de renegociação de sua dívida, firmado com a cobrança, seu contador da quantidade de dias de atraso é congelado, somente sendo reativado quando este associado atrasar uma das parcelas de seu acordo, por um período maior que o do contador; desta forma, caso incluíssemos este público no estudo, misturaríamos duas populações muito distintas. Com isso, um modelo estatístico específico para clientes nesta situação será desenvolvido e implementado na companhia até meados do ano seguinte.

4.3 Aplicação da Metodologia Estatística

O trabalho aqui desenvolvido apresenta de um lado uma variável resposta Y que assume resultados entre 0 e 1, os quais correspondem respectivamente às contas com altíssima probabilidade de terem seus saldos contabilizados como prejuízo (ou seja, atingirem 180 dias em atraso) e às contas com altíssima probabilidade de terem os financiamentos de seus saldos negociados ou mesmo terem seus débitos liquidados de uma só vez; e, de outro lado, um conjunto de variáveis independentes X que serão expostas mais adiante (Seção 4.3.1).

Colocar-se-ão em prática, a partir de agora, as técnicas estatísticas descritas no Capítulo 3 para construirmos a equação do modelo de *Charge Off Scoring*.

4.3.1 *Extração das Informações*

As variáveis que podem auxiliar no conhecimento do cliente, ao longo de seu relacionamento com a instituição enquanto inadimplente, são diversas, contudo, diante da atual situação de migração de sistemas na qual se encontra a organização, podemos dispor apenas de algumas delas.

Isto posto, do ponto de vista da administração de crédito, destacam-se as seguintes variáveis independentes para compor o modelo:

Risco do associado se tornar inadimplente

O risco de um associado se tornar inadimplente é medido através de uma pontuação conhecida como *behavior score* na empresa na qual foi desenvolvido este trabalho. Trata-se de uma variável discreta que resume muito bem o comportamento do cliente ao longo do seu relacionamento com a instituição estudada. Ele corresponde a uma nota dada para cada conta do cliente que define seu risco de inadimplência. Assim, a conta pode ser de altíssimo risco (E), de alto risco (D), de médio risco (C), de baixo risco (B) e de baixíssimo risco (A). Existe também a possibilidade de que não haja tempo de relacionamento suficiente com o cliente para que seja definido seu risco de inadimplência, o qual será classificado como Z, ou seja, risco indefinido. Esta classificação também é dada às contas canceladas tanto por inadimplência quanto por desistência do associado.

Obviamente, espera-se que contas com alto risco de inadimplência tenham maior probabilidade de não efetuarem pagamento, até que estas atinjam o atraso suficiente para serem contabilizadas como prejuízo, se comparadas às de menor risco. Aliás, algumas empresas no Brasil utilizam esta variável para priorizar todas as suas ações de cobrança.

Entretanto, o *behavior scoring* é melhor utilizado para prevenir a inadimplência, na medida em que evidencia quem são os clientes passíveis de um acompanhamento mais próximo, que pode ser feito, por exemplo, pela diminuição dos limites de crédito

ou até mesmo pelo bloqueio preventivo dos cartões, mediante algum sinal de maior risco.

Por outro lado, dado que já ocorreu a inadimplência, acredita-se que a predição de pagamento anterior à contabilização da dívida como prejuízo, pode ser maximizada pela presença de outras variáveis além da informação do *behavior score*.

A expectativa é que o *behavior score* seja uma variável bastante significativa no modelo proposto uma vez que ele caracteriza o comportamento de pagamentos do cliente ao longo de seu relacionamento com a instituição. Porém, espera-se que o modelo torne-se mais significativo estatisticamente com a presença de outras variáveis não contempladas no modelo de *behavior scoring*, permitindo, desta forma, maior separação dos clientes que poderão pagar em até três meses daqueles que não pagarão. Esta variável foi observada nos meses T e T-1 para cada conta estudada.

Saldo Devedor do Associado

Esta é a primeira das variáveis contínuas empregadas no modelo. O valor do saldo devido pelo associado foi observado nos meses T (mês em que o associado está com sua fatura mensal atrasada entre 91 e 120 dias) e T-1.

A experiência mostrava que saldos devedores maiores deveriam gerar maiores dificuldades para pagar. Por esta razão, a priorização das contas para cobrança era feita pelo saldo devedor, ou seja, saldos maiores eram cobrados primeiro.

Limite de Crédito

Todo usuário de cartão de crédito tem um valor máximo para a concessão do crédito que não pode ser ultrapassado. Caso ocorram algumas despesas superiores a este limite estabelecido (*overlimit*), aqui incluídas também as compras parceladas, o associado tem seu cartão bloqueado. O limite de crédito é mais uma das variáveis contínuas consideradas como relevantes para o desenvolvimento do modelo.

Percentual de utilização do limite de crédito

A relação entre o saldo devedor e o limite de crédito do cliente inadimplente compuseram outra variável contínua que será considerada: o percentual de utilização do limite nos meses T e T-1.

Espera-se que quanto maior o percentual utilizado do limite de crédito, menor a probabilidade de quitação da dívida, uma vez que o limite de crédito está relacionado à renda do cliente. Se o percentual de utilização é alto, acredita-se que o pagamento do saldo devedor pode comprometer a renda mensal do associado e, por essa razão, este pagamento tem menor chance de ocorrer.

Tempo de relacionamento

Uma das características mais importantes no processo de gerenciamento de contas a receber de uma empresa de cartão de crédito é conhecer o tempo de relacionamento com o cliente. Espera-se que clientes inadimplentes que tiveram um período grande de relacionamento com a empresa, estejam passando por dificuldades temporárias e que, por essa razão, tenham maior probabilidade de pagamento, antes de que suas dívidas sejam contabilizadas pela empresa como prejuízo.

Alto tempo de relacionamento significa comportamento anterior com pagamentos em dia, pois, caso contrário, a conta do cliente estaria cancelada por inadimplência (quando atinge 91 dias em atraso).

Valor da Multa/Mora pelo atraso no pagamento da fatura mensal

Assim que o associado atrasa o pagamento de sua fatura mensal, incorrem sobre o seu valor devido multa e mora de 2% e 1% respectivamente. Os valores desta variável somente foram considerados neste estudo para o mês T-1, e, para justificar esta decisão, apresentamos a Figura 4.3:

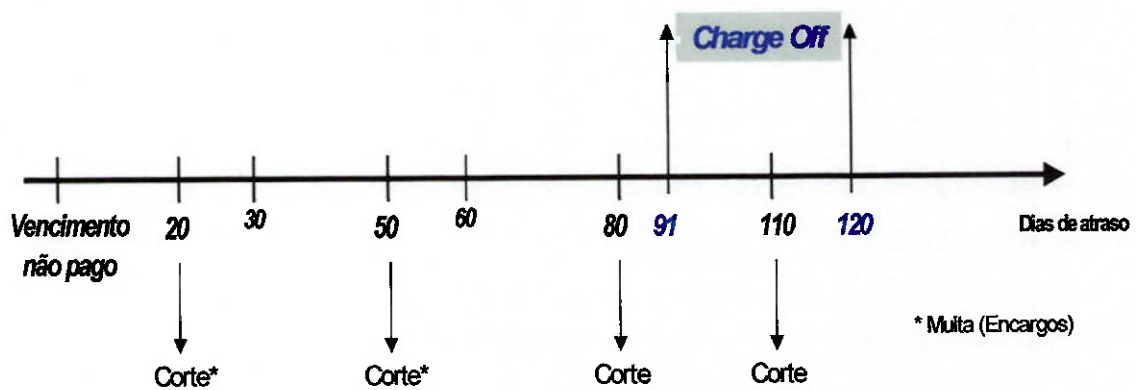


Figura 4.3 – Linha do Tempo da cobrança de valores da variável multa/mora

A data em que são contabilizadas todas as transações financeiras realizadas naquele mês para cada conta é conhecida como a data do corte da fatura mensal (que corresponde a dez dias antes da data do vencimento). Diante disso, para o público selecionado para a construção do modelo (com atraso entre 91 e 120 dias), podemos identificar duas situações distintas para esta variável:

- caso a conta selecionada para a amostra esteja entre 91 e 110 dias de atraso, o mês atual (T) será o do corte no dia 80 e o mês anterior (T-1) será o do corte no dia 50 e;
- caso a conta selecionada para a amostra esteja entre 111 e 120 dias de atraso, o mês atual (T) será o do corte no dia 110 e o mês anterior (T-1) será o do corte no dia 80.

Assim, o valor desta variável em T corresponderá a zero em qualquer das situações descritas acima (pois multa/mora somente são cobradas nos cortes dos dias 20 e 50) e, por esta razão, não faria sentido considerarmos seu valor na construção do modelo.

Desta forma, resulta apenas o valor da multa/mora pelo atraso no pagamento da fatura mensal em T-1 para ser considerado na confecção do modelo.

Bandeira do produto do associado

As bandeiras foram apresentadas na Seção 2.1.2 deste trabalho, assim, neste momento, cumpre-nos apenas destacar que a instituição na qual o modelo foi desenvolvido, trabalha com apenas duas bandeiras: a Visa e a Mastercard.

Produto do associado

Por motivos de confidencialidade, não serão explicitados quais os diversos tipos de cartões de crédito (produtos) dos quais dispõe a administradora na qual se baseou o estudo. Assim sendo, simplesmente salientaremos que existe uma família de produtos com preços e benefícios diferenciados para cada segmento do mercado, alta, média ou baixa renda; os segmentos alvo de mercado dessa empresa são os de média e alta renda, enquanto que o público de baixa renda ocupa menor destaque em termos de importância na definição de estratégias de atuação. Esta variável pode apresentar 21 valores distintos, ou seja 21 produtos, os quais designamos pelas letras A até U.

Local de Residência do Associado

Aqui representado por regiões de código de endereçamento postal (CEP), o local de residência do associado é outra variável que foi considerada como relevante, na identificação dos associados que têm maior probabilidade de não quitarem suas dívidas, antes que seus saldos sejam contabilizados como prejuízo. Estas regiões de CEP foram compostas, de acordo com a proximidade das localidades, de modo a permitir maior clareza nas análises posteriores, e são apresentadas a seguir:

- Aracaju – SE
- Belém – PA
- Centro de São Paulo – SP
- Curitiba – PR
- Estado de Santa Catarina
- Estado do Espírito Santo
- Estado do Maranhão
- Estado do Piauí
- Estados da Paraíba e Rio Grande do Norte
- Fortaleza – CE

- Grande São Paulo
- Interior da Bahia
- Interior de Minas Gerais
- Interior do Ceará
- Interior do Pará
- Interior do Paraná
- Interior do Pernambuco
- Interior do Rio Grande do Sul
- Litoral do Estado de São Paulo
- Porto Alegre - RS
- Recife - PE
- Região Centro-oeste
- Região de Araçatuba - SP
- Região de Bauru - SP
- Região de Belo Horizonte - MG
- Região de Campinas - SP
- Região de Campos - RJ
- Região de Duque de Caxias - RJ
- Região de Niterói - RJ
- Região de Nova Iguaçu - RJ
- Região de Presidente Prudente - SP
- Região de Ribeirão Preto - SP
- Região de São José do Rio Preto - SP
- Região de Sorocaba - SP
- Região de Taubaté - SP
- Região de Volta Redonda - RJ
- Região do ABCD-SP
- Região Norte

- Salvador - BA
- Centro do Rio de Janeiro - RJ
- Zona Leste de São Paulo - SP
- Zona Norte de São Paulo - SP
- Zona Norte do Rio de Janeiro - RJ
- Zona Oeste de São Paulo - SP
- Zona Oeste do Rio de Janeiro - RJ
- Zona Sul de São Paulo - SP

4.3.2 *Exploração das Informações*

Inicialmente, convém ressaltar que atribuímos uma sigla mnemônica a cada variável descrita na Seção 4.3.1, conforme apresentado na Tabela 4.2:

Variável	Sigla
Valor da Multa pelo atraso no pagamento da fatura mensal em T-1	MULTA_T-1
Saldo devedor em T-1	SALDO_T-1
Saldo devedor em T	SALDO_T
Limite de Crédito	LIM_T
Percentual de utilização do limite de crédito em T-1	UTLIM_T-1
Percentual de utilização do limite de crédito em T	UTLIM_T
Tempo de Relacionamento	IDADE
Local de residência do associado	CD_CEP
Risco do associado se tornar inadimplente em T-1	BEHAV_T-1
Risco do associado se tornar inadimplente em T	BEHAV_T
Produto do associado	INSTIT
Bandeira do produto do associado	BANDEIRA

Tabela 4.2 – Siglas Mnemônicas das Variáveis

Isto posto, após a definição e coleta da amostra, e antes de partir para a construção do modelo propriamente dito, um passo importante foi a análise exploratória dos dados.

Realizou-se, então, a sumarização dos aspectos importantes do conjunto de dados. A Tabela 4.3 exhibe os dados condensados de localização e dispersão para cada variável independente contínua, os quais incluem a tendência central dos dados (valor médio da variável), o valores mínimo e máximo, a mediana, a moda, a amplitude, o desvio padrão e, finalmente, a variância amostral.

Variável	Média	Mediana	Moda	Menor	Maior	Amplitude	Desvpad ³	Variância
MULTA_T-1	29,25	18,01	0,00	0,00	413,33	413,33	41,83	1.749,42
SALDO_T-1	1.379,40	930,67	10,15	5,22	13.986,42	13.981,20	1.522,39	2.317.668,35
SALDO_T	1.605,30	1.079,99	144,75	100,14	14.847,52	14.747,38	1.777,15	3.158.265,39
LIM_T	1.401,22	1.000,00	600,00	70,00	11.250,00	11.180,00	1.259,60	1.586.602,03
UTLIM_T-1	1,05	1,09	0,37	0,00	8,23	8,23	0,61	0,37
UTLIM_T	1,22	1,26	1,39	0,01	9,79	9,79	0,71	0,51
IDADE	18,85	10,83	4,67	3,33	222,67	219,33	24,86	617,88

Tabela 4.3 – Estatísticas Descritivas da base de dados de desenvolvimento

Através destas estatísticas, pode-se observar a presença de *outliers* (valores extremos, muito distantes da média dos demais) e as razões para isso foram prontamente identificadas: a base de dados de onde a amostra foi extraída possuía alguns registros com ausência de valores para algumas variáveis, bem como identificou-se registros com valores diferentes dos reais (saldos cujo valor estava multiplicado por 100, por exemplo).

Da amostra de 17.013 contas, 170 apresentaram valores de variáveis fora do padrão, sendo 118 contas “más” e 52 contas “boas”. Percebe-se, então, que a exclusão destas em nada afetará a proporção entre contas “boas” e “más” da amostra; com isso, chegou-se à composição definitiva da amostra de acordo com a Tabela 4.4:

Bases de Dados	Contas “Más”		Contas “Boas”		Total	
	N	%	N	%	N	%
Base de desenvolvimento	11.571	68,70%	5.272	31,30%	16.843	100%
Base de validação	2.191	67,44%	1.058	32,56%	3.249	100%
Total	13.762	68,49%	6.330	31,51%	20.092	100%

Tabela 4.4 – Composição definitiva das bases de dados amostrais

³ Desvpad corresponde ao desvio padrão.

4.3.3 Matriz de Correlação

Uma análise importante nessa fase é a relação entre as variáveis explicativas, com o objetivo de se conhecer a multicolinearidade entre elas, ou seja, se a presença de uma das variáveis no modelo acarretará a eliminação de outra. Se a colinearidade entre duas variáveis for alta, e ambas forem correlacionadas com a variável resposta, a presença de apenas uma delas será suficiente para explicar parte da variação da variável resposta. A análise de multicolinearidade permite prever, portanto, quais variáveis não deverão ter presença conjunta no modelo de *charge off scoring*.

Assim, foi construída uma matriz que visa apontar as correlações entre as variáveis independentes contínuas.

Considerou-se, conforme citado na Seção 3.2, que valores superiores a $|0,6|$ são indicadores de que as variáveis envolvidas estejam altamente correlacionadas. A matriz de correlação exibida na Tabela 4.5 apresenta as diversas variáveis contínuas e suas inter-relações.

VARIÁVEIS	SALDO_T-1	SALDO_T	LIM_T	UTLIM_T-1	UTLIM_T	MULTA_T-1	IDADE
SALDO_T-1	100%						
SALDO_T	69%	100%					
LIM_T	60%	60%	100%				
UTLIM_T-1	45%	43%	-12%	100%			
UTLIM_T	43%	45%	-12%	68%	100%		
MULTA_T-1	61%	64%	55%	27%	37%	100%	
IDADE	20%	19%	25%	1%	-1%	14%	100%

Tabela 4.5 – Matriz de correlação das variáveis contínuas

Desta forma, analisando-se a matriz anterior, pode-se notar (valores em vermelho) que os pares de variáveis relacionados abaixo, apresentam alto índice de correlação (superior a $|0,6|$):

- SALDO_T \Leftrightarrow SALDO_T-1
- LIM_T \Leftrightarrow SALDO_T-1
- LIM_T \Leftrightarrow SALDO_T
- UTLIM_T \Leftrightarrow UTLIM_T-1
- MULTA_T-1 \Leftrightarrow SALDO_T-1
- MULTA_T-1 \Leftrightarrow SALDO_T

Com isso, seria possível excluir algumas variáveis da Regressão Logística; contudo, devido à mais significativa destas correlações exceder em 9% o mínimo valor definido para que uma correlação seja considerada alta, optou-se por manter todas as variáveis contínuas, deixando esta tarefa para ser realizada, caso seja necessária, pelo procedimento de *Stepwise* da Regressão Logística.

4.3.4 Aplicação da Análise Bivariada

Nesta etapa do trabalho, é feita a análise da relação entre os possíveis valores das variáveis discretas explicativas com a variável resposta. Por exemplo, confronta-se a informação do produto que o cliente dispõe com o fato do cliente ter pago ou não, antes que sua dívida tivesse sido contabilizada como prejuízo. Dependendo do grau e da forma desta relação, avalia-se se a presença daquela variável no modelo pode ser esperada ou não; e se essa variável deve ser categorizada (agrupada em poucas categorias) ou utilizada da forma como foi coletada, respectivamente.

Conforme mencionado, analisou-se aqui somente as variáveis independentes discretas candidatas a compor o modelo. Para cada variável, construímos uma tabela que, além de apresentar as quantidades de contas “boas” e “más” por valor da variável explicativa, exhibe uma coluna que expõe os agrupamentos resultantes, dada a

semelhança entre as razões de *Odds* (razão entre a probabilidade de um evento ocorrer e não ocorrer, no caso a quitação da dívida antes da contabilização como prejuízo).

Risco de Inadimplir em T

A Tabela 4.6 apresenta a análise bivariada do risco do associado inadimplir em T pela variável resposta.

BEHAV_T	"Maus"		"Bons"		Total		Odds	Grupo
	N	%	N	%	N	%		
A	0	-	0	-	0	-	-	-
B	8	57,14%	6	42,86%	14	100,00%	0,75	1
C	619	61,17%	393	38,83%	1.012	100,00%	0,63	
D	3.537	66,36%	1.793	33,64%	5.330	100,00%	0,51	2
E	6.858	69,62%	2.993	30,38%	9.851	100,00%	0,44	3
Z	549	86,32%	87	13,68%	636	100,00%	0,16	4
Total	11.571	68,70%	5.272	31,30%	16.843	100,00%	0,46	-

Tabela 4.6 - Análise Bivariada do Risco de Inadimplir em T x Variável Resposta

Pode-se perceber que a razão de *Odds* para a amostra completa é de 0,46 (linha de total da Tabela 4.6). Assim, nota-se que os valores da variável independente com significância estatística são aqueles cuja razão de *Odds* está distante de 0,46, tanto para mais (caso do *behavior score* B, cujo percentual de contas “boas” é maior que o percentual de contas “boas” de toda a amostra), quanto para menos (caso do *behavior score* E, cujo percentual de contas “boas” é menor que o percentual de contas “boas” de toda a amostra).

Dada a semelhança entre as razões de *Odds*, agrupou-se alguns dos valores das variáveis discretas, de modo a podermos criar um número menor de variáveis *dummy* a serem consideradas na modelagem. Assim, foi criado o grupo formado pelo *behavior score* B e C (união também motivada pela pouca quantidade de observações em B) e os grupos 2, 3 e 4 (*behavior score* D, E e Z, respectivamente).

Contudo, faz-se necessária a aplicação de um teste estatístico que compare as proporções de contas “boas” dos valores da variável explicativa agrupados, de maneira a verificarmos se a diferença entre eles não é estatisticamente significativa, justificando, assim, sua união em um mesmo grupo; compara-se os valores observados e esperados. O teste estatístico que faz esta comparação é o teste de homogeneidade (também conhecido como teste qui-quadrado). A estatística qui-quadrado é uma medida do quão distantes estão os valores observados dos valores esperados. Sua fórmula é

$$X^2 = \sum \frac{(\text{valor_observado} - \text{valor_esperado})^2}{\text{valor_esperado}}$$

e sua soma é sobre todos as células das

linhas x colunas da tabela dos dados.

Desta forma, aplicando-se este teste para o grupo formado pelo *behavior score* B e C temos (Tabela 4.7):

<i>Behavior</i> <i>Score</i>	Valores observados			Valores esperados	
	"Maus"	"Bons"	Total	"Maus"	"Bons"
B	8	6	14	8,56	5,44
C	619	393	1.012	618,44	393,56
Total	627	399	1.026	-	-

Tabela 4.7 – Valores observados e esperados para *Behavior Score* B e C⁴

A partir dos dados da Tabela 4.7 pode-se calcular o valor da estatística X^2 que corresponde a $0,018038 + 0,0283447 + 0,000124 + 0,0001947 = 0,046701$. Sabendo-se que tem-se 1 grau de liberdade (número de colunas - 1 x número de linhas - 1), obtém-se da tabela de valores críticos da distribuição qui-quadrada do Anexo 2 o valor de 3,84 que é maior que o encontrado (0,046701). Com isso, pode-se afirmar, com 95% de confiança que os valores B e C desta variável são homogêneos.

⁴ Os valores esperados são calculados para a célula c_{ij} multiplicando-se o total da coluna i pelo total da linha j e, em seguida, dividindo-se esse resultado pelo total de dados da tabela.

Contudo, deve-se ainda identificar se os quatro grupos criados não são homogêneos entre si. Para tanto, deve-se aplicar o teste de Qui-quadrado sobre eles, somando-se as quantidades de contas “boas” e “más” dos integrantes de cada grupo, conforme segue:

Grupo	Valores observados			Valores esperados	
	"Maus"	"Bons"	Total	"Maus"	"Bons"
1	627	399	1.026	704,85	321,15
2	3.537	1.793	5.330	3.661,67	1.668,33
3	6.858	2.993	9.851	6.767,55	3.083,45
4	549	87	636	436,93	199,07
Total	11.571	5.272	16.843	-	-

Tabela 4.8 – Valores observados e esperados para os Grupos de *Behavior Score*

Com os dados da Tabela 4.8 calcula-se o valor da estatística X^2 que corresponde a $8,60 + 4,24 + 1,25 + 28,75 + 18,87 + 9,32 + 2,65 + 63,09 = 136,74$. Com 3 graus de liberdade, verifica-se na tabela de valores críticos da distribuição qui-quadrada do Anexo 2 o valor de 7,81 que é menor que o encontrado (136,74). Com isso, pode-se afirmar, com 95% de confiança que os grupos da variável BEHAV_T não são homogêneos.

Assim como o ocorrido com a matriz de correlação, embora tenhamos identificado grupos cuja Razão de *Odds* aponte para pouca distinção entre contas “boas” e “más”, não serão, em princípio, excluídos da regressão. Com isso, serão criadas variáveis *dummy* para todos os grupos de 1 a 4: *behav_t_1*, *behav_t_2*, *behav_t_3* e *behav_t_4*, , respectivamente.

Risco de Inadimplir em T-1

Para a variável do risco de inadimplir em T-1 não foi possível criar grupos com razão de *Odds* semelhante dada a diferença de comportamento entre todos os valores que esta variável pode assumir. Ainda assim, verificou-se se cada categoria desta variável difere das demais aplicando-se o teste de Qui-quadrado. Seu resultado aponta para ausência de homogeneidade pois o valor obtido para a estatística foi 110,65, maior que o valor crítico de 9,49 (4 graus de liberdade).

Contudo, pode-se identificar, através da Tabela 4.9, as classificações B, C, E e Z como as mais significantes para esta variável (razões de *Odds* distantes de 0,46); no entanto, todos os grupos de 1 a 5 novamente foram convertidos, respectivamente, na variáveis *dummy*: behav_t1_1, behav_t1_2, behav_t1_3, behav_t1_4 e behav_t1_5.

BEHAV_T-1	"Maus"		"Bons"		Total		Odds	Grupo
	N	%	N	%	N	%		
A	0	-	0	-	0	-	-	-
B	138	58,97%	96	41,03%	234	100,00%	0,70	1
C	1.622	63,11%	948	36,89%	2.570	100,00%	0,58	2
D	3.604	67,45%	1.739	32,55%	5.343	100,00%	0,48	3
E	5.233	72,62%	1.973	27,38%	7.206	100,00%	0,38	4
Z	974	65,37%	516	34,63%	1.490	100,00%	0,53	5
Total	11.571	68,70%	5.272	31,30%	16.843	100,00%	0,46	-

Tabela 4.9 - Análise Bivariada do Risco de Inadimplir em T-1 x Variável Resposta

Produto

Quanto ao tipo de produto do associado, pode-se notar com a análise da Tabela 4.10, que grupos de razão de *Odds* semelhante puderam ser criados; contudo, nenhum destes apresentou significativa relevância estatística. Ainda assim, criamos três variáveis *dummy* para representá-lo na regressão logística que define o modelo, são eles: instit_1, instit_2 e instit_3.

INSTIT	"Maus"		"Bons"		Total		Odds	Grupo
	N	%	N	%	N	%		
A	345	72,78%	129	27,22%	474	100,00%	0,37	1
B	1.286	72,61%	485	27,39%	1.771	100,00%	0,38	
C	1.187	72,64%	447	27,36%	1.634	100,00%	0,38	
D	211	68,51%	97	31,49%	308	100,00%	0,46	2
E	224	68,50%	103	31,50%	327	100,00%	0,46	
F	202	68,71%	92	31,29%	294	100,00%	0,46	
G	202	68,47%	93	31,53%	295	100,00%	0,46	
H	492	68,24%	229	31,76%	721	100,00%	0,47	
I	375	67,81%	178	32,19%	553	100,00%	0,47	
J	285	67,86%	135	32,14%	420	100,00%	0,47	
K	274	67,99%	129	32,01%	403	100,00%	0,47	
L	597	67,84%	283	32,16%	880	100,00%	0,47	
M	544	67,83%	258	32,17%	802	100,00%	0,47	
N	966	67,84%	458	32,16%	1.424	100,00%	0,47	
O	219	67,80%	104	32,20%	323	100,00%	0,47	
P	2.151	67,81%	1.021	32,19%	3.172	100,00%	0,47	
Q	685	67,82%	325	32,18%	1.010	100,00%	0,47	
R	724	65,58%	380	34,42%	1.104	100,00%	0,52	3
S	209	64,91%	113	35,09%	322	100,00%	0,54	
T	215	64,95%	116	35,05%	331	100,00%	0,54	
U	178	64,73%	97	35,27%	275	100,00%	0,54	
Total	11.571	89,25%	5.272	40,67%	12.964	100,00%	0,46	-

Tabela 4.10 - Análise Bivariada do Produto x Variável Resposta

Dada a semelhança entre as razões de *Odds* dos produtos da Tabela 4.10, foram criados 3 grupos para a variável produto do associado. Estes grupos foram submetidos ao teste de homogeneidade e os resultados obtidos foram os seguintes:

- Grupo 1: com 2 graus de liberdade e 95% de confiança tem-se o valor crítico de 5,99 (Anexo 2), que quando comparado ao valor da estatística $X^2 = 0,00549$ do grupo, aponta para a existência de homogeneidade.
- Grupo 2: com 13 graus de liberdade e 95% de confiança encontra-se o valor crítico de 22,36 (Anexo 2), que comparado ao valor da estatística $X^2 = 0,2948$ do grupo, atesta a existência de homogeneidade.

- Grupo 3: com 3 graus de liberdade e 95% de confiança obtém-se o valor crítico de 7,81 (Anexo 2), que comparado ao valor da estatística $X^2 = 0,1155$ do grupo, garante, também, a homogeneidade deste grupo.

É importante ressaltar que os grupos não são homogêneos entre si, pois a estatística $X^2 = 42,2706$ é superior ao valor crítico de 5,99 (2 graus de liberdade).

Bandeira

Assim como a variável anterior, percebe-se (Tabela 4.11) que a bandeira do produto do associado não corresponde a uma informação significativa para colaborar na separação de bons e maus do modelo; contudo, foram criadas variáveis *dummy* para mais esta variável explicativa: bandeira_1 e bandeira_2.

BANDEIRA	"Maus"		"Bons"		Total		Odds	Grupo
	N	%	N	%	N	%		
I	5.102	67,76%	2.427	32,24%	7.529	100,00%	0,48	1
II	6.469	69,45%	2.845	30,55%	9.314	100,00%	0,44	2
Total	11.571	68,70%	5.272	31,30%	16.843	100,00%	0,46	-

Tabela 4.11 - Análise Bivariada da Bandeira x Variável Resposta⁵

O teste de homogeneidade aplicado às duas opções desta variável retornou o valor 5,52955 para estatística X^2 que, quando comparado com o valor de 3,84 da tabela do Anexo 2, garante com 95% de confiança a não homogeneidade das bandeiras.

⁵ Por motivo de confidencialidade da informação, foram omitidas as quantidades de cartões de cada bandeira.

Local de Residência do Cliente

Quanto à região de residência do cliente, a associação destas em grupos através da razão de *Odds* foi importante para viabilizar o tratamento de um conjunto de dados muito abrangente; a grande quantidade de valores que esta variável pode assumir torna esta prática fundamental.

Mais uma vez, pode-se perceber que existem diversos valores desta variável cujas razões de *Odds* apresentam semelhança (Tabela 4.12). Desta forma, 11 grupos distintos foram criados, sendo apenas 10 destes relevantes estatisticamente e, por consequência, com maiores chances de “entrar” no modelo. Assim sendo, as seguintes variáveis *dummy* foram definidas para representar a região de residência do associado: gr_1, gr_2, gr_3, gr_4, gr_5, gr_6, gr_7, gr_8, gr_9, gr_10 e gr_11.

CD_CEP	"Maus"		"Bons"		Total		Odds	Grupo
	N	%	N	%	N	%		
Aracaju - SE	125	80,13%	31	19,87%	156	0,93%	0,25	1
Interior do Pará	20	80,00%	5	20,00%	25	0,15%	0,25	
Região de Nova Iguaçu - RJ	288	80,22%	71	19,78%	359	2,13%	0,25	
Região de Duque de Caxias - RJ	441	77,78%	126	22,22%	567	3,37%	0,29	2
Zona Norte do Rio de Janeiro - RJ	92	77,31%	27	22,69%	119	0,71%	0,29	
Interior da Bahia	1.633	77,25%	481	22,75%	2.114	12,55%	0,29	
Estado do Maranhão	193	77,51%	56	22,49%	249	1,48%	0,29	3
Região de Volta Redonda - RJ	86	74,78%	29	25,22%	115	0,68%	0,34	
Recife - PE	66	74,16%	23	25,84%	89	0,53%	0,35	
Estado do Espírito Santo	127	74,27%	44	25,73%	171	1,02%	0,35	4
Zona Oeste do Rio de Janeiro - RJ	18	72,00%	7	28,00%	25	0,15%	0,39	
Interior do Paraná	263	72,05%	102	27,95%	365	2,17%	0,39	
Centro do Rio de Janeiro - RJ	698	72,03%	271	27,97%	969	5,75%	0,39	5
Região de Niterói - RJ	994	71,77%	391	28,23%	1.385	8,22%	0,39	
Interior do Ceará	560	71,70%	221	28,30%	781	4,64%	0,39	
Fortaleza - CE	147	71,71%	58	28,29%	205	1,22%	0,39	6
Salvador - BA	187	69,78%	81	30,22%	268	1,59%	0,43	
Interior do Rio Grande do Sul	294	69,34%	130	30,66%	424	2,52%	0,44	
Porto Alegre - RS	292	68,38%	135	31,62%	427	2,54%	0,46	7
Estado do Piauí	37	68,52%	17	31,48%	54	0,32%	0,46	
Região de Presidente Prudente - SP	35	68,63%	16	31,37%	51	0,30%	0,46	
Curitiba - PR	241	67,70%	115	32,30%	356	2,11%	0,48	8
Grande São Paulo	470	65,55%	247	34,45%	717	4,26%	0,53	
Interior do Pernambuco	198	65,56%	104	34,44%	302	1,79%	0,53	
Interior de Minas Gerais	430	64,66%	235	35,34%	665	3,95%	0,55	9
Zona Oeste de São Paulo - SP	334	64,60%	183	35,40%	517	3,07%	0,55	
Estados da Paraíba e Rio Grande do Norte	110	64,71%	60	35,29%	170	1,01%	0,55	
Região de Ribeirão Preto - SP	149	64,22%	83	35,78%	232	1,38%	0,56	10
Centro de São Paulo - SP	198	64,29%	110	35,71%	308	1,83%	0,56	
Zona Leste de São Paulo - SP	373	64,20%	208	35,80%	581	3,45%	0,56	
Litoral do Estado de São Paulo	180	64,29%	100	35,71%	280	1,66%	0,56	11
Região de São José do Rio Preto - SP	75	64,10%	42	35,90%	117	0,69%	0,56	
Região de Belo Horizonte - MG	393	63,08%	230	36,92%	623	3,70%	0,59	
Região de Bauru - SP	66	62,86%	39	37,14%	105	0,62%	0,59	12
Região de Campos - RJ	94	61,44%	59	38,56%	153	0,91%	0,63	
Zona Sul de São Paulo - SP	388	61,20%	246	38,80%	634	3,76%	0,63	
Zona Norte de São Paulo - SP	239	59,75%	161	40,25%	400	2,37%	0,67	13
Região de Campinas - SP	367	59,77%	247	40,23%	614	3,65%	0,67	
Região Centro-oeste	72	60,00%	48	40,00%	120	0,71%	0,67	
Região Norte	75	58,14%	54	41,86%	129	0,77%	0,72	14
Região de Sorocaba - SP	105	58,01%	76	41,99%	181	1,07%	0,72	
Região de Taubaté - SP	127	58,26%	91	41,74%	218	1,29%	0,72	
Região do ABCD-SP	18	58,06%	13	41,94%	31	0,18%	0,72	15
Estado de Santa Catarina	86	57,72%	63	42,28%	149	0,88%	0,73	
Belém - PA	165	57,89%	120	42,11%	285	1,69%	0,73	
Região de Araçatuba - SP	22	57,89%	16	42,11%	38	0,23%	0,73	16
Total	11.571	68,70%	5.272	31,30%	16.843	100,00%	0,46	

Tabela 4.12 - Análise Bivariada do Local de Residência x Variável Resposta

Estes grupos foram também submetidos ao teste de homogeneidade e o resultados obtidos foram os seguintes:

- Grupo 1: 2 graus de liberdade, 95% de confiança, valor crítico = 5,99 (Anexo 2), $X^2 = 0,0012 \rightarrow$ grupo homogêneo
- Grupo 2: 3 graus de liberdade, 95% de confiança, valor crítico = 7,81 (Anexo 2), $X^2 = 0,0752 \rightarrow$ grupo homogêneo
- Grupo 3: 2 graus de liberdade, 95% de confiança, valor crítico = 5,99 (Anexo 2), $X^2 = 0,3270 \rightarrow$ grupo homogêneo
- Grupo 4: 5 graus de liberdade, 95% de confiança, valor crítico = 11,07 (Anexo 2), $X^2 = 0,0388 \rightarrow$ grupo homogêneo
- Grupo 5: 5 graus de liberdade, 95% de confiança, valor crítico = 11,07 (Anexo 2), $X^2 = 0,4118 \rightarrow$ grupo homogêneo
- Grupo 6: 1 grau de liberdade, 95% de confiança, valor crítico = 3,84 (Anexo 2), $X^2 = 0,6608 \rightarrow$ grupo homogêneo
- Grupo 7: 7 graus de liberdade, 95% de confiança, valor crítico = 14,07 (Anexo 2), $X^2 = 0,0565 \rightarrow$ grupo homogêneo
- Grupo 8: 1 grau de liberdade, 95% de confiança, valor crítico = 3,84 (Anexo 2), $X^2 = 0,2342 \rightarrow$ grupo homogêneo
- Grupo 9: 1 grau de liberdade, 95% de confiança, valor crítico = 3,84 (Anexo 2), $X^2 = 1,4519 \rightarrow$ grupo homogêneo
- Grupo 10: 2 graus de liberdade, 95% de confiança, valor crítico = 5,99 (Anexo 2), $X^2 = 3,1234 \rightarrow$ grupo homogêneo
- Grupo 11: 6 graus de liberdade, 95% de confiança, valor crítico = 12,59 (Anexo 2), $X^2 = 0,0133 \rightarrow$ grupo homogêneo

É importante ressaltar que os grupos não são homogêneos entre si, pois a estatística $X^2 = 319,5596$ é superior ao valor crítico de 18,31 (10 graus de liberdade).

4.3.5 Variáveis Dummy

Diante das conclusões extraídas da Seção 4.3.4, as quais que proporcionaram a formação de grupos para cada variável discreta, resumiu-se as variáveis *dummy* que serão utilizadas na Regressão Logística através da Tabela 4.13:

Variável Independente	Variável <i>Dummy</i>
BEHAV_T-1	beh t1 1
	beh t1 2
	beh t1 3
	beh t1 4
	beh t1 5
BEHAV_T	beh t 1
	beh t 2
	beh t 3
	beh t 4
CD_CEP	gr 1
	gr 2
	gr 3
	gr 4
	gr 5
	gr 6
	gr 7
	gr 8
	gr 9
	gr 10
	gr 11
INSTIT	instit 1
	instit 2
	instit 3
BANDEIRA	bandeira 1
	bandeira 2

Tabela 4.13 – Conversão das variáveis independentes discretas em variáveis *dummy*

Com isso, definimos foram estipuladas todas as variáveis que serão utilizadas na Regressão Logística (Tabela 4.14), que será detalhada na Seção 4.3.6.

Variável	Tipo	Variável	Tipo
beh t1_1	Dummy	beh t_1	Dummy
beh t1_2	Dummy	beh t_2	Dummy
beh t1_3	Dummy	beh t_3	Dummy
beh t1_4	Dummy	beh t_4	Dummy
beh t1_5	Dummy	instit_1	Dummy
gr_1	Dummy	instit_2	Dummy
gr_2	Dummy	instit_3	Dummy
gr_3	Dummy	bandeira_1	Dummy
gr_4	Dummy	bandeira_2	Dummy
gr_5	Dummy	SALDO T-1	Contínua
gr_6	Dummy	SALDO T	Contínua
gr_7	Dummy	LIM T	Contínua
gr_8	Dummy	UTLIM T-1	Contínua
gr_9	Dummy	UTLIM T	Contínua
gr_10	Dummy	MULTA T-1	Contínua
gr_11	Dummy	IDADE	Contínua

Tabela 4.14 – Variáveis finais para entrada no modelo

4.3.6 Aplicação da Regressão Logística

De posse das variáveis exibidas na Tabela 4.14, aplicou-se a regressão logística através da metodologia *stepwise* descrita por Hosmer e Lemeshow (1989).

Cumpramos ressaltar que neste trabalho será adotada significância de 0,05. Isto posto, o *Passo 0* foi utilizado para definir o valor da constante, de acordo como se segue:

- *Passo 0*: Definiu-se o valor da constante do modelo igual a 0,1367, caso em que todas as contas analisadas teriam a mesma probabilidade $\pi = e^{0,1367} / (1 + e^{0,1367}) = 53,41\%$ de não serem contabilizadas como prejuízo e calculou-se o valor de seu logaritmo de verossimilhança: $L_0 = -7.308,544$.

- *Passo 1:* Adicionou-se a variável $MULTA_T-1$ e calculou-se o novo valor do logaritmo de verossimilhança $L_1 = -7.154,614$. Em seguida, obteve-se o valor da estatística G:

$$G = -2 (L_0 - L_1) = -2 \cdot (-7.308,544 + 7.154,614) = 307,8600$$

Consultando as tabelas da X^2 (Qui-quadrado) percebe-se que a probabilidade associada a este valor com 1 grau de liberdade é inferior a 0,0005, ou seja, menor que o nível de significância especificado de 0,05; portanto a variável $MULTA_T-1$ entra no modelo.

- *Passo 2:* Realizando o mesmo procedimento com a adição da variável $SALDO_T$ obtém-se o seguinte valor para o logaritmo de verossimilhança $L_2 = -6.878,9125$.

Calculando o valor da estatística G:

$$G = -2 (L_1 - L_2) = -2 \cdot (-7.154,614 + 6.878,9125) = 551,4030$$

Verificando na tabela da Qui-quadrado (X^2) percebe-se que a probabilidade é ainda inferior a 0,0005 (1 grau de liberdade), que por sua vez é menor que 0,05; logo a variável entra no modelo.

Testa-se, ainda, a permanência da variável $MULTA_T-1$:

$$G' = -2 (L_{1(SALDO_T)} - L_2) = -2 \cdot (-7.213,144 + 6.878,9125) = 674,463$$

Este valor resulta em uma probabilidade é inferior a 0,0005 (1 grau de liberdade), que por sua vez é menor que 0,05. A variável $MULTA_T-1$ é, então, mantida no modelo.

- *Passo 3:* Devido aos próximos passos do algoritmo serem análogos, serão suprimidas as próximas etapas, exibindo apenas os resultados finais. A Figura 4.4 ilustra o processo de regressão ora aplicado.

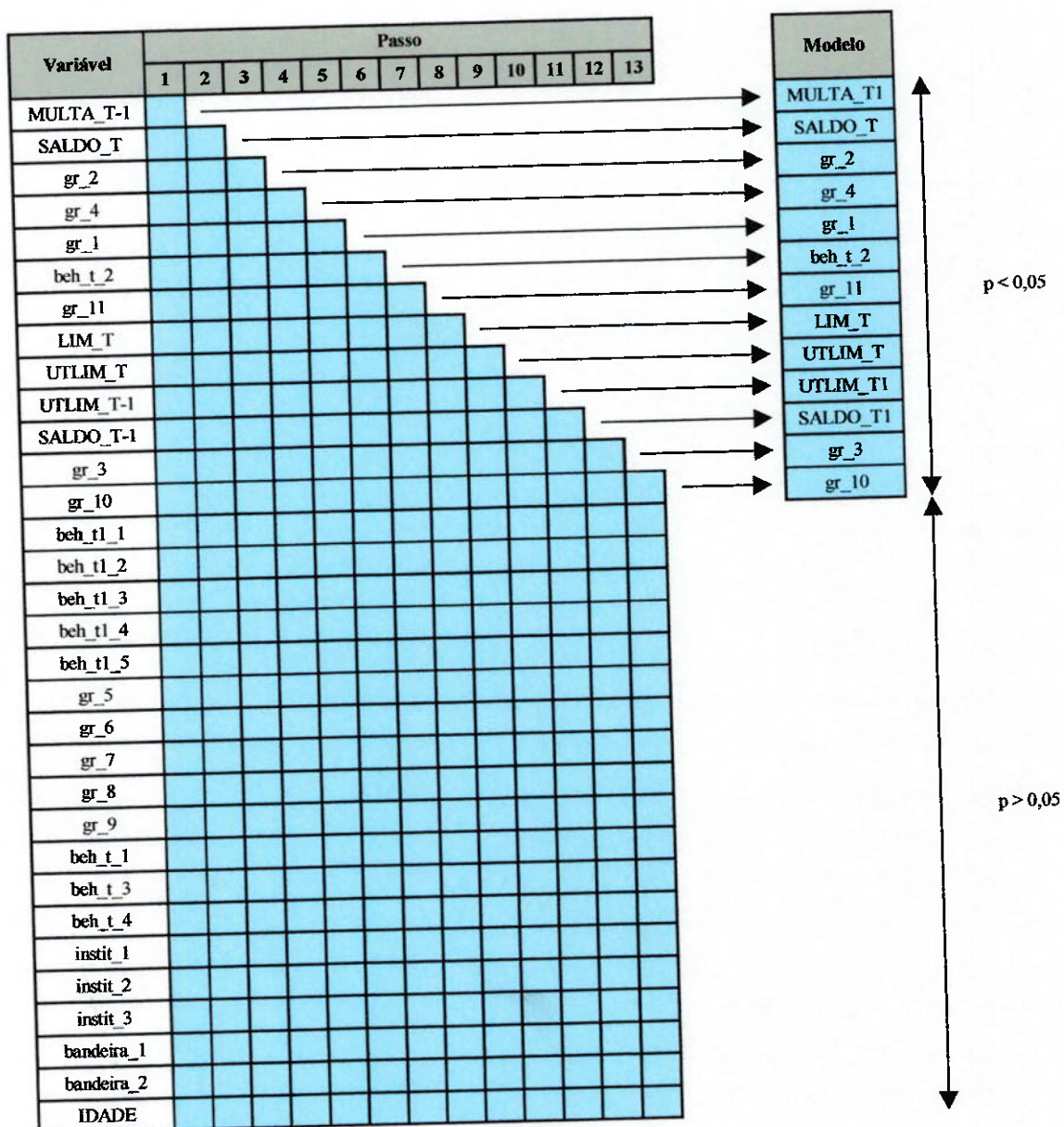


Figura 4.4 – Método *Stepwise* de seleção das variáveis

Os valores da estatística G para as demais etapas podem ser visualizados na Tabela 4.15.

Passo	Variável de Entrada		G	P
	Incluída	Removida		
1	MULTA_T-1	-	307,8600	< 0,0001
2	SALDO_T	-	551,4030	< 0,0001
3	gr_2	-	65,4040	< 0,0001
4	gr_4	-	35,3919	< 0,0001
5	gr_1	-	38,7768	< 0,0001
6	beh_t_4	-	23,2845	< 0,0001
7	gr_11	-	12,3447	0,0004
8	LIM_T	-	9,7940	0,0018
9	UTLIM_T	-	6,3481	0,0118
10	UTLIM_T-1	-	29,0731	< 0,0001
11	SALDO_T-1	-	52,3007	< 0,0001
12	gr_3	-	6,6645	0,0098
13	gr_10	-	5,1056	0,0238

Tabela 4.15 – Sumário do Método de Seleção *Stepwise*

A análise da entrada de uma variável no modelo é determinada por um acréscimo marginal de significância ao modelo medido pela Função Verossimilhança. Esse acréscimo proporciona uma variação na probabilidade do coeficiente da variável candidata a entrar no modelo ser zero. Até que esse acresce ultrapasse a probabilidade fixada (5%), as variáveis entram no modelo.

A modelagem através da técnica de regressão logística, que incluiu as variáveis SALDO_T-1, SALDO_T, LIM_T, UTLIM_T-1, UTLIM_T, MULTA_T-1, beh_t_4, gr_1, gr_2, gr_3, gr_4, gr_10 e gr_11, apresentou os seguintes valores de coeficientes (Tabela 4.16):

Variável de Entrada	Coefficiente	Erro Padrão	Qui-quadrado	P
Constante	0,1367	0,0691	3,9071	0,0481
SALDO_T-1	0,0009	0,0001	42,3999	< 0,0001
SALDO_T	-0,0014	0,0001	137,5667	< 0,0001
LIM_T	0,0001	0,0000	8,4722	0,0036
UTLIM_T-1	-1,9939	0,2525	62,3745	< 0,0001
UTLIM_T	1,7587	0,2186	64,7295	< 0,0001
MULTA_T-1	0,0289	0,0015	385,4625	< 0,0001
beh_t_4	-0,6315	0,1348	21,9485	< 0,0001
gr_1	-0,7320	0,1261	33,7066	< 0,0001
gr_2	-0,5315	0,0621	73,3438	< 0,0001
gr_3	-0,3551	0,1444	6,0424	0,0140
gr_4	-0,2998	0,0526	32,4574	< 0,0001
gr_10	0,2628	0,1166	5,0849	0,0241
gr_11	0,4396	0,1251	12,3362	0,0004

Tabela 4.16 – Análise dos Coeficientes de Máxima Verossimilhança

4.3.7 Resultados

Da Tabela 4.16 da Seção 4.3.6, extraiu-se os valores do vetor A, o vetor de coeficientes das variáveis na equação da Regressão Logística:

$$g(x) = a_0 + a_1X_1 + a_2X_2 + \dots + a_kX_k = \ln \left[\frac{\pi}{1 - \pi} \right]$$

De acordo com os coeficientes obtidos (Tabela 4.16), chegamos à seguinte equação:

$$g(x) = 0,1367 + 0,0009.SALDO_T-1 - 0,0014.SALDO_T + 0,0001.LIM_T - 1,9939.UTLIM_T-1 + 1,7587.UTLIM_T + 0,0289.MULTA_T-1 - 0,6315.beh_T4 - 0,7320.gr_1 - 0,5315.gr_2 - 0,3551.gr_3 - 0,2998.gr_4 + 0,2628.gr_10 + 0,4396.gr_11$$

A Tabela 4.17 mostra todos os parâmetros do modelo final obtido através do método de regressão logística.

Atributo	Pontuação	Atributo	Pontuação
Constante	0,1367	Multa em T-1	
Saldo Devedor em T		MULTA_T1	0,0289
SALDO_T	-0,0014	Risco de inadimplir	
Saldo Devedor em T		Indefinido (Z)	-0,6315
SALDO_T1	0,0009	Local de Residência	
Utilização do Limite em T		Aracaju-SE, Interior do Pará, Região de Nova Iguaçu-RJ	-0,7320
UTLIM_T	1,7587	Região de Duque de Caxias-RJ, Zona Norte do Rio de Janeiro-RJ, Interior da Bahia e Maranhão	-0,5315
Utilização do Limite em T-1		Região de Volta Redonda-RJ, Recife-PE e Espírito Santo	-0,3551
UTLIM_T1	-1,9939	Zona Oeste e Centro do Rio de Janeiro-RJ, Região de Niterói-RJ, Interior do Ceará, Interior do Paraná e Fortaleza	-0,2998
Limite de Crédito		Zona Norte de São Paulo-SP, Região de Campinas-SP e Região Centro-oeste	0,2628
LIM_T	0,0001	Região Norte, Região de Sorocaba-SP, Região de Taubaté-SP, Região do ABCD-SP, Santa Catarina, Belém-PA e Região de Araçatuba-SP	0,4396

Tabela 4.17 – Tabela de Parâmetros (Resultados do Modelo)

As variáveis contínuas estão representadas na Tabela 4.17 em azul, as variáveis *dummy* em verde e a constante em vermelho. Os parâmetros, representados pelos valores dentro da tabela, indicam o quanto uma variável influenciou na probabilidade de quitação da dívida. Portanto, a presença de uma variável com sinal positivo do parâmetro significa que ela aumenta a probabilidade de que aquela conta não tenha seu saldo contabilizado como prejuízo. A intensidade deste aumento corresponde ao próprio parâmetro.

Para se obter a probabilidade de quitação do saldo devedor antes que este seja contabilizado como prejuízo, deve-se transformar a equação segundo os critérios apresentados na Seção 3.4.2. A equação que resulta é:

$$\pi(x) = \frac{e^{0,1367 + 0,0009 \cdot \text{SAL DO}_T - 1 - 0,0014 \cdot \text{SAL DO}_T + 0,0001 \cdot \text{LIM}_T - 1,9939 \cdot \text{UTL IM}_T - 1 + 1,7587 \cdot \text{UTL IM}_T + 0,0289 \cdot \text{MUL TA}_T - 1 - 0,6315 \cdot \text{beh}_T - 0,7320 \cdot \text{gr}_1 - 0,5315 \cdot \text{gr}_2 - 0,3551 \cdot \text{gr}_3 - 0,2998 \cdot \text{gr}_4 + 0,2628 \cdot \text{gr}_10 + 0,4396 \cdot \text{gr}_11}}{1 + e^{0,1367 + 0,0009 \cdot \text{SAL DO}_T - 1 - 0,0014 \cdot \text{SAL DO}_T + 0,0001 \cdot \text{LIM}_T - 1,9939 \cdot \text{UTL IM}_T - 1 + 1,7587 \cdot \text{UTL IM}_T + 0,0289 \cdot \text{MUL TA}_T - 1 - 0,6315 \cdot \text{beh}_T - 0,7320 \cdot \text{gr}_1 - 0,5315 \cdot \text{gr}_2 - 0,3551 \cdot \text{gr}_3 - 0,2998 \cdot \text{gr}_4 + 0,2628 \cdot \text{gr}_10 + 0,4396 \cdot \text{gr}_11}}$$

5 ANÁLISE DOS RESULTADOS

Este capítulo objetiva analisar os resultados obtidos, apresentar algumas informações relevantes sobre a validação e manutenção do modelo, bem como a aplicação do teste de Kolmogorov-Smirnov para verificar seu percentual de separação. Além disso, os resultados obtidos com o modelo serviram de base para a definição de algumas estratégias de cobrança, as quais foram testadas através da realização de um estudo piloto.

5.1 Análise das variáveis que compuseram o modelo

Com o objetivo de fornecer interpretações financeiramente coerentes para o que foi determinado pelo modelo, apresentamos, a seguir, a análise por variável compositora da equação de $\pi(x)$:

Risco do associado se tornar inadimplente

Conforme já mencionado, o *behavior score* é uma nota dada a cada conta do cliente, a qual define seu risco de inadimplência. Diante desta nota, a conta do associado pode ser classificada como de baixíssimo risco (A), de baixo risco (B), de médio risco (C), de alto risco (D), de altíssimo risco (E) ou de risco indefinido (Z). Dentre estas opções, somente entrou no modelo a classificação Z. A variável *dummy* de *behavior score* Z (*beh_t_4*) entrou com valor de parâmetro negativo (-0,6315).

Uma vez não havendo conta com *behavior score* A dentro da população estudada (pois é praticamente impossível haver alguma conta com baixíssimo risco de inadimplir, já inadimplente a mais de 90 dias), era de se esperar, portanto, que os *behavior scores* B e C, de menores riscos no caso, fossem incluídos no modelo com valor de parâmetro associado positivo, ou mesmo se ausentassem, que foi exatamente o ocorrido. Situação análoga era esperada para os riscos altíssimo e alto de inadimplir (E e D

respectivamente): a expectativa era de que estes não apresentassem alta significância neste momento da inadimplência do cliente (91 a 120 em atraso), porém, caso fossem incluídos, esperava-se que aparecessem com valor negativo, ou seja, diminuindo a probabilidade de quitação do saldo. Ao mesmo tempo, o risco indefinido (Z), por ser composto, na totalidade da amostra de estudo, por contas já canceladas (contas que não despertam mais interesse da instituição), apresentou parâmetro com valor negativo no modelo. Os resultados são, portanto, coerentes com a expectativa.

Local de Residência do Associado

Esta variável apresentou as 6 regiões nas quais residem os associados que apresentam relevância para o modelo, são elas: as regiões 1 a 4 que apresentaram os parâmetros -0,7320, -0,5315, -0,3551 e -0,2998 respectivamente, e as regiões 10 e 11 as quais influenciam positivamente a probabilidade de regularização do saldo devedor (na ordem +0,2628 e +0,4396). Cumpre ressaltar que as 4 primeiras regiões contém os estados do nordeste e um dos estados do Sudeste, os quais historicamente apresentam elevados índices de inadimplência.

Limite de Crédito

Esta variável reflete o valor máximo do crédito concedido pela instituição ao associado, incluindo aqui compras parceladas. O limite de crédito foi selecionado para entrar no modelo, e seu coeficiente apresentou valor positivo: 0,0001. Diante disso, pode-se perceber que quanto maior o limite de crédito, maior será a probabilidade do cliente quitar sua dívida antes que esta possa ser considerada prejuízo (variável resposta próxima de 1).

Este resultado surpreende um pouco. Era esperado que limites de crédito mais altos não afetassem positivamente a probabilidade de quitação da dívida, pois a hipótese é de que estes limites implicariam em maiores saldos devedores, apresentando, por consequência, maior dificuldade de regularização; entretanto, o que obteve-se como

resultado é justamente o contrário, embora o valor do parâmetro para a variável limite seja pequeno, quando comparado aos coeficientes das variáveis de saldo (SALDO_T e SALDO_T1).

Saldo Devedor em T e T-1

Essa variável representa o valor da dívida do cliente no período entre os 91 e 120 dias em atraso (T) e entre 61 e 90 (T-1). Analisando-se ambas as variáveis isoladamente, percebe-se, então, que o modelo proposto apresenta coeficientes antagônicos para os dois meses do histórico, ou seja, as contas com maiores saldos devedores em T-1 e menores saldos devedores em T serão aquelas com maior probabilidade de quitação da dívida. Mantendo-se esta análise isolada, o resultado parece incoerente para T-1, uma vez que a hipótese é de que saldos devedores maiores aumentariam sua dificuldade de regularização.

Contudo, sabendo-se que estas variáveis apresentaram correlação significativa (69%, segundo Tabela 4.5), o que justifica-se pelo fato de comumente haver pouca alteração no valor devido pelo inadimplente quando sua dívida passa da faixa de 61-90 para 91-120 dias em atraso, deve-se, então, analisá-las em conjunto. Esta análise, nos remete a um resultado perfeitamente coerente: aquelas contas cujo saldo devido foi reduzido de um mês para o outro, ou seja, maior saldo em T-1 e menor saldo em T serão as contas com maior probabilidade de regularização do mesmo, uma vez que algum pagamento foi realizado neste intervalo de tempo.

Por outro lado, aquelas contas em que houve crescimento no saldo de um mês para o outro (parcelas de compras feitas no passado sendo postadas na conta neste momento, por exemplo), assim como as contas cujo saldo manteve-se estável terão menores probabilidades de quitação, uma vez que estes associados não fizeram nenhum pagamento no período, o que é perfeitamente coerente.

Utilização do Limite de Crédito em T e T-1

Essa variável representa o quanto o cliente efetivou de compras ou saques, diante do limite de crédito concedido pela instituição para ele (tanto em T quanto em T-1). Novamente, analisando-se ambas as variáveis isoladamente, percebe-se, então, que o modelo proposto apresenta coeficientes antagônicos para os dois meses do histórico, ou seja, as contas com menor utilização do limite em T-1 e maior utilização em T serão aquelas com maior probabilidade de quitação da dívida. Mantendo-se esta análise isolada, o resultado parece incoerente para T, pois espera-se que quanto mais o cliente utilizar de seu limite, menor a chance de regularização de seus débitos.

Porém, sabendo-se que estas variáveis possuem correlação significativa (68%, de acordo com os dados da Tabela 4.5), o que justifica-se pelo fato de raramente haver alteração no índice de utilização do limite no período 61 a 120 dias em atraso (T-1 e T), estas serão analisadas em conjunto. Esta análise confirma a hipótese de um resultado surpreendente: aquelas contas cuja utilização do limite cresceu de um mês para o outro serão as contas com maior probabilidade de regularização do mesmo. Em paralelo a isto, aquelas contas em que houve decréscimo na utilização, assim como as contas cuja utilização manteve-se estável terão menores probabilidades de quitação (não há surpresa apenas para esta última situação, pois para a utilização ter se mantido estável, o associado não deve ter apresentado nenhum pagamento que reduzisse seu saldo, e por consequência, sua utilização do limite no período).

Valor da Multa/Mora pelo atraso no pagamento da fatura mensal

Assim que o associado atrasa o pagamento de sua fatura mensal, são cobrados dele dois valores percentuais de seu saldo devido, referentes a multa e a mora (2% e 1% respectivamente). Conforme mencionado e justificado na Seção 4.3.1, somente foi considerado o valor de multa/mora em T-1, o qual apresentou relevância estatística, entrando no modelo com o coeficiente positivo 0,0289.

Esse resultado é consistente com o esperado: sabendo-se que as contas que apresentam valor de multa/mora nulo em T-1 são aquelas com atraso superior a 111 dias, esperava-se que estas apresentassem maior risco de atingir os 180 dias em atraso (prejuízo); paralelamente a isto, as contas que apresentaram valor diferente de zero estarão necessariamente com atraso inferior a 110 dias e, por conseqüência, devem apresentar maior probabilidade de quitação da dívida (Figura 4.3). Contudo, nota-se que as contas cujo valor da multa é mais alto, receberão um aumento na probabilidade de ocorrência do evento 1 (quitação), o que não necessariamente é verdade, uma vez que a multa/mora corresponde a um percentual do saldo, e esperava-se que maiores saldo implicassem em maiores dificuldades para regularização.

5.2 *Ponto de Corte*

Diante dos resultados alcançados, apresentados na Seção 4.3.7, decidiu-se criar faixas de pontuações de maneira a classificar as contas dos associados em cinco níveis de risco distintos, de acordo com o grau de risco desta conta atingir os 180 dias em atraso, e ter, por conseqüência, seu saldo contabilizado como prejuízo. Esta classificação visa facilitar o entendimento dos resultados que serão discutidos neste capítulo.

Cumpre salientar que, a fim de melhorar a compreensão dos resultados, a probabilidade da conta ter seu saldo renegociado antes que este seja contabilizado como prejuízo foi multiplicada por 100 e denominada de pontuação.

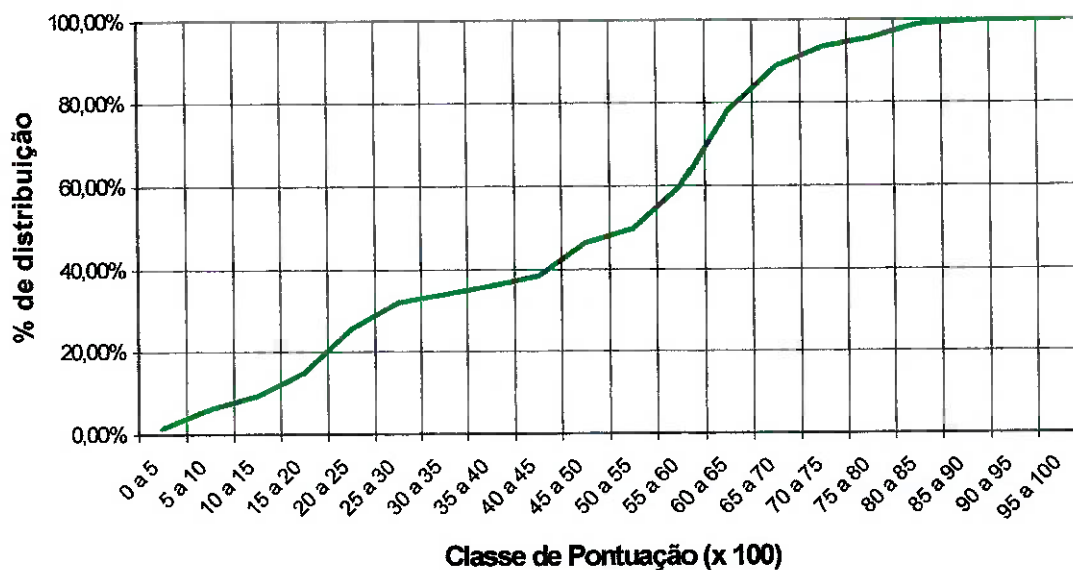


Figura 5.1 – Distribuição de frequência acumulada de toda a população por pontuação

Dada a distribuição de frequência da Figura 5.1, realizou-se o corte da pontuação de acordo com cada quintil¹. Assim, cinco faixas de pontuação foram caracterizadas, permitindo a análise de seu respectivo percentual de contas recuperadas (mais uma vez: entende-se por conta recuperada aquela que tem seu saldo renegociado com a instituição, através do pagamento da primeira parcela de um acordo firmado, seja esta parcela única ou não), as quais podem ser verificadas Tabela 5.1:

¹ O quintil é qualquer das separatrizes que dividem a população de uma distribuição de frequência em quantidade iguais a múltiplos inteiros de um quinto da população.

Pontuação (x 100)	Maus	Bons	Total	% Bons % Maus	% Contas Recuperadas
Até 20	309	3.060	3.369	0,10	9,17%
de 21 a 45	510	2.859	3.369	0,18	15,14%
de 46 a 57	1.039	2.330	3.369	0,45	30,84%
de 58 a 63	1.614	1.754	3.368	0,92	47,92%
Mais de 64	1.800	1.568	3.368	1,15	53,44%
Total	5.272	11.571	16.843	0,46	31,30%

Tabela 5.1 – Faixas de pontuação de “bons” e “maus”

Através da Tabela 5.1, percebe-se a acentuada diferença entre os percentuais de recuperação das duas primeiras faixas e pontuação e das duas últimas, estando a faixa central com percentual muito parecido com o da população como um todo. Diante disso, tornou-se possível classificar as pontuações nas cinco categorias, conforme segue (Tabela 5.2):

Classificação	Risco do saldo ser contabilizado como prejuízo	Pontuação
COFF 0	Altíssimo	000 a 020
COFF 1	Alto	021 a 045
COFF 2	Médio	046 a 057
COFF3	Baixo	058 a 063
COFF 4	Baixíssimo	064 a 100

Tabela 5.2 – Classificação das Pontuações (x 100) do Modelo

5.3 Análise de Desempenho

A adequação deste modelo foi através da realização do teste de Kolmogorov-Smirnov (KS), o qual revela se as distribuições de contas “boas” e “más” são iguais ou diferentes. A estatística KS é a maior diferença ordenada observada entre duas funções de distribuição empíricas. De acordo com Caouette; Altman; Narayanan (1999), caso este valor seja próximo de zero, diz-se que as distribuições estão superpostas; caso esteja entre 30 e 60%, considera-se como duas distribuições com bom índice de separação; e finalmente, se o KS for próximo de 100%, afirma-se que as distribuições são completamente disjuntas (Figura 5.2). Portanto, quanto maior o valor de KS, maior é a eficiência do modelo em segregar duas distribuições, que, no caso, são as contas “boas” das contas “más”.

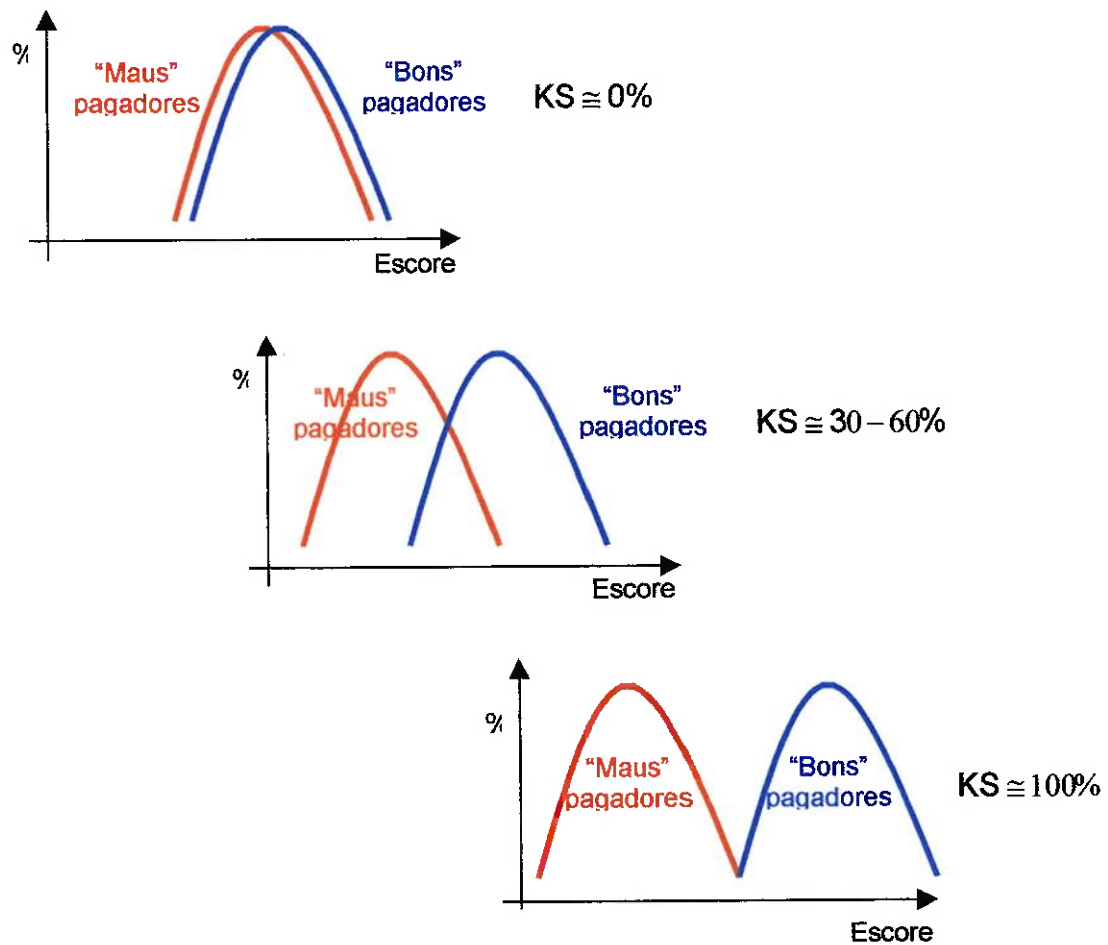


Figura 5.2 – Teste de adequação de Kolmogorov-Smirnov – Adaptado de Souza (2000)

O teste de Kolmogorov-Smirnov (estatística KS), aplicado através da ferramenta SAS, mostra que a separação deste modelo é de 35%. Esse resultado pode ser considerado bom diante da atual disponibilidade de variáveis para a confecção do modelo, uma vez que os modelos dessa natureza tidos como viáveis devem ter separação entre 30 e 60%. Segundo Cordell et al (1998), é bastante comum encontrar bons modelos com separação de 30 a 50%.

A Figura 5.3 mostra a distribuição de freqüência dos dois grupos de contas ("boas" e "más") analisadas por faixa de pontuação definidas na Seção 5.2. Percebe-se que a maior concentração de contas "más" encontra-se nas primeiras faixas de

pontuação (de 000 a 040), enquanto que a maior proporção de contas “boas” encontra-se nas faixas superiores (de 060 a 100).

Percebe-se, entretanto, que existe uma região de sobreposição dos dois grupos (de 040 a 060) em que é difícil definir se uma conta pertence a um grupo ou outro. Este resultado mostra que o modelo não fornece uma total separação dos dois grupos.

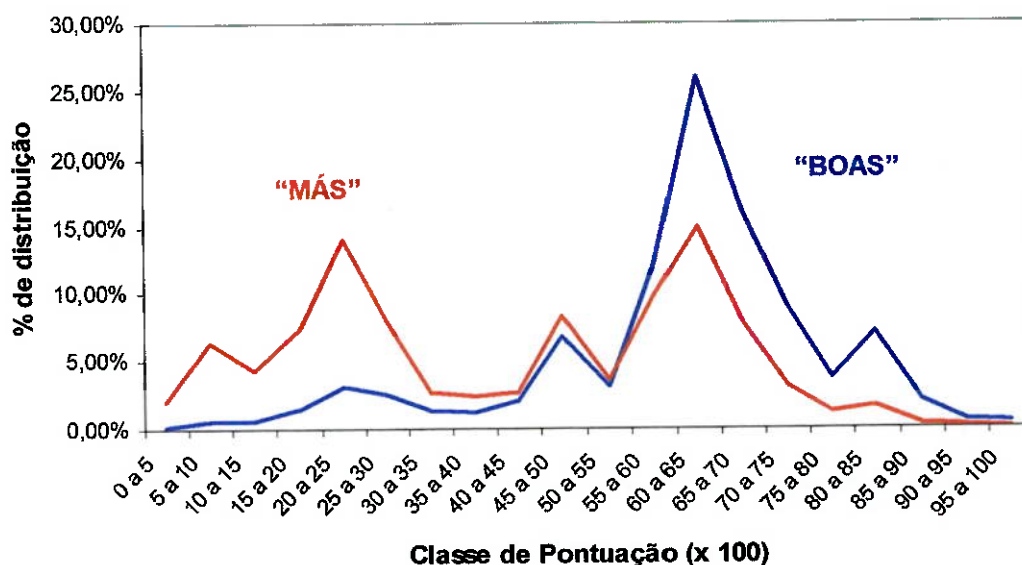


Figura 5.3 – Distribuição dos bons e maus clientes por pontuação

Ao se analisar a Figura 5.4, é mais fácil verificar o grau de separação do modelo. Esse gráfico mostra a distribuição de frequência acumulada por faixa de pontuação dos dois grupos analisados. Percebe-se que até a pontuação 025, a proporção de contas “más” é bem maior do que a de “boas”. Até a pontuação 040, embora a proporção de contas “boas” cresça, ela não supera 10% (pontuação menor ou igual a 040). Da mesma maneira, pontuação superior a 070 tem pouco mais de 10% de contas “más”.

Portanto, para as faixas de escore extremas, o modelo fornece ótima separação entre as contas “boas” e “más”. A região central, entretanto, é que não permite identificar com grande precisão a qual grupo pertenceria uma determinada conta. Fazê-lo, teria uma probabilidade de erro significativa.

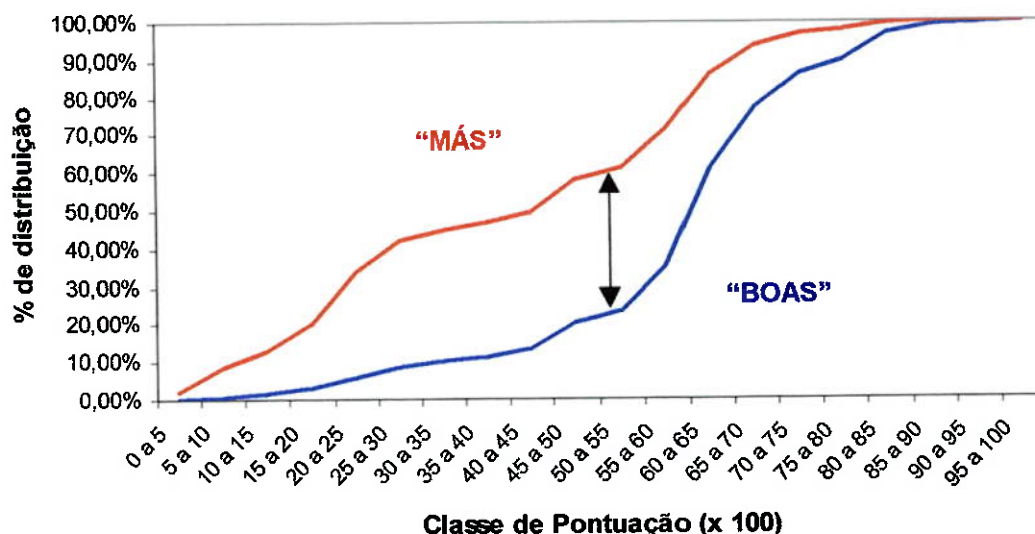


Figura 5.4 – Distribuição acumulada dos bons e maus clientes por pontuação

5.4 Validação

É muito importante que antes da implementação final do modelo seja realizado um teste numa amostra de contas inadimplentes, que não foram utilizadas durante o processo de desenvolvimento do modelo, para assegurar que o mesmo e suas predições sejam compatíveis com os resultados obtidos no desenvolvimento e com a realidade dessas contas. Caso haja diferenças significativas, as causas devem ser encontradas e resolvidas antes da implementação em ambiente de produção.

Para comprovar a validade deste modelo, inicialmente estabeleceu-se o seguinte critério: se as contas “boas” foram classificadas como de altíssimo e alto risco de prejuízo, assumiu-se que foram classificadas incorretamente. Da mesma forma, se uma conta “má” foi designada como de baixo ou baixíssimo risco, também foi considerado erro.

Isto posto, a Tabela 5.3 mostra a distribuição das contas “boas” e “más” por classificação de risco de não renegociação da dívida. Observa-se que 29% $((1.754+1.568)/11.571)$ das contas “más” foram classificadas como “boas de baixo ou

baixíssimo risco) e 16% $((309+510)/5.272)$ das contas “boas” foram classificadas com “más”. Portanto, 25% das 16.843 contas foram classificadas erroneamente.

Pontuação	Risco de Prejuízo	Classificação	“Boas”	“Más”	Total
COFF 0	Altíssimo	“Más”	309	3.060	3.369
COFF 1	Alto		510	2.859	3.369
COFF 2	Médio	Incertas	1.039	2.330	3.369
COFF 3	Baixo	“Boas”	1.614	1.754	3.368
COFF 4	Baixíssimo		1.800	1.568	3.368
Total	-	-	5.272	11.571	16.843

Tabela 5.3 – Distribuição dos “bons” e “maus” por score

A zona de sobreposição é a região compreendida pela classificação COFF 2 (de médio risco de prejuízo ou incerta), que representa 20% $(3.369/16.843)$ da amostra. Nela, não é possível decidir se uma conta é “boa” ou “má”, pois 20% das contas más e 20% das contas boas se encontram com essa classificação.

Ainda com respeito a esta validação, foram dadas notas a uma base de 3.249 contas para validação apresentada na Tabela 4.4 (1.058 contas “boas” e 2.191 contas “más”). Os resultados dos parâmetros obtidos com a base de desenvolvimento foram atribuídos à base de validação, a qual retornou, quando submetida ao teste de Kolmogorov-Smirnov, um índice de separação de 33,5%, o que é extremamente próximo ao alcançado com a base de desenvolvimento.

5.5 Implementação

Mais importante do que desenvolver este modelo de *charge off scoring* é implementá-lo. De nada adianta possuir um modelo com ótimo poder de predição e

separação se este não é aplicável. A maneira como utilizá-lo é um pensamento que deve estar presente durante todo o processo de planejamento para que, no momento de implementá-lo, as principais dificuldades inerentes a este processo estejam previstas e resolvidas ou, ao menos, minimizadas.

Ao se implementar o modelo de *charge off scoring*, deve-se ter a certeza de que o processo operacional está preparado para utilizá-lo. Isto significa que as pessoas que irão utilizar esta ferramenta precisam conhecer as suas funcionalidades e objetivos e, mais do que isso, que estejam treinadas para utilizá-lo e gerenciá-lo. Além disso, sistemas que sustentam a utilização e manutenção do modelo devem ser instalados e verificados periodicamente no que se refere ao cálculo das pontuações e registro das ações tomadas com cada cliente.

Para automatizar e tornar sistêmico a atribuição de pontuação às contas contidas dentro do intervalo de atraso de 91 a 120, foi desenvolvido um programa em SAS que extrai todos os dados do *Data Warehouse* da instituição, constitui uma tabela com todas as informações sobre as variáveis que foram selecionadas para compor o modelo, cria as variáveis *dummy* necessárias e, finalmente, executa a atribuição de pontuação a esta base utilizando-se, para tanto, da Tabela 4.17 de parâmetros do capítulo anterior.

5.5.1 Estudo piloto: Definição de estratégias e resultados

Um aspecto importante para a implementação do modelo de *charge off scoring* é que as políticas e estratégias de cobrança devem estar alinhadas com a utilização deste modelo. Isso significa que, para cada faixa de score do modelo, é necessário que seja planejada uma estratégia diferenciada, de acordo com a probabilidade do cliente inadimplente atingir os 180 dias em atraso.

Assim, percebe-se que um dos objetivos da construção do modelo de *charge off scoring* é identificar segmentos em que se possa traçar estratégias de cobrança diferenciadas. O resultado da Tabela 5.3, revela que podem existir pelo menos três estratégias: uma para as classificações COFF 0 e COFF 1 (altíssimo e alto risco de

prejuízo), que apontam as contas verdadeiramente “más”, outra para as classificações COFF 3 e COFF 4 (baixo e baixíssimo risco de prejuízo), que indicam as “melhores” contas e, finalmente, outra estratégia para a classificação COFF 2 (médio risco), que, apesar de representar uma parcela pequena da população, mistura contas “boas” e “más”.

Estas estratégias foram definidas e adotadas (Tabela 5.4), para um estudo piloto, com o objetivo de validar ainda mais o modelo proposto. Foram separadas 5.021 contas na faixa de atraso de 90 a 120 dias, no mês de julho de 2.001. Estas foram pontuadas de acordo com o modelo proposto (utilizando-se, para tanto de suas informações nos meses de junho e julho de 2.001), e classificadas em cada uma das cinco categorias de risco de contabilização do saldo como prejuízo. Em seguida, cada conta foi submetida à estratégia de cobrança adotada específica para sua classificação de risco.

RISCO DE PREJUÍZO	AÇÃO DE COBRANÇA
Altíssimo	Envio das contas para escritórios externos de cobrança
Alto	
Médio	Cobrança diferenciada na Central Telefônica de Recuperação
Baixo	Cobrança comum através do Segmento especialista no nível de Risco E (90 - 120) na Central Telefônica de Recuperação
Baixíssimo	

Tabela 5.4 – Estratégias adotadas para o estudo piloto

O resultado descrito na Tabela 5.5 seguinte permite analisar a eficiência do modelo de *charge off scoring* para a empresa estudada. Percebe-se que das 2.006 contas classificadas como de altíssimo e alto risco, e nas quais foram aplicadas estratégias mais “agressivas” de cobrança (atuação de escritórios externos de cobrança especialistas na recuperação de contas já contabilizadas como prejuízo), 30% delas renegociaram seu

saldo devido através de acordos. Por outro lado, das 2.061 contas que foram classificadas como sendo de baixo e baixíssimo risco, 51% renegociaram suas dívidas.

Risco de Prejuízo	Renegociação	Prejuízo	Total
Altíssimo / Alto	589	1.417	2.006
Médio	346	608	954
Baixo / Baixíssimo	1.051	1.010	2.061
Total	1.986	3.035	5.021

Tabela 5.5 – Resultado das contas pontuadas após implementação das estratégias

Outro ponto a ser destacado é que se fosse aplicada a mesma estratégia para todos os níveis de risco, era de se esperar que a taxa de contabilização como prejuízo após 3 meses fosse muito maior para os níveis de risco alto e altíssimo do que para os níveis baixo e baixíssimo.

Entretanto, com a aplicação de estratégias diferenciadas para cada nível de risco, o percentual de contabilização destas contas como prejuízo aproximou-se muito mais para os níveis de risco nos dois extremos (alto e baixo risco). Enquanto para as contas de altíssimo e alto risco o percentual de prejuízo foi de 71%, para as contas de baixo e baixíssimo risco, o percentual foi de 49%. Isso mostra que a estratégia adotada para as contas com alto risco estão sendo eficazes no que se refere a evitar que os saldos fossem lançados a prejuízo e, por consequência, o aumento no nível de rentabilização de toda a carteira de inadimplentes.

Ressalta-se, ainda, a confirmação de que o nível de risco médio (COFF 2) concentra-se realmente numa zona de sobreposição, apesar do emprego de uma estratégia de cobrança diferenciada na central telefônica ter proporcionado resultados que apontaram mais contas com resposta positiva (renegociação). Das 954 contas que receberam essa classificação, 36% renegociaram seus saldos; em paralelo a isso, as demais 64% atingiram 180 dias em atraso. Percebe-se, portanto, que trata-se de uma

situação em que o modelo proposto realmente não conseguiu prever se essas contas pertencem ao grupo de alto ou de baixo risco, e, por isso, receberam a classificação de médio risco. Embora o resultado esteja bastante coerente com o modelo – pois não se esperava das contas classificadas como de médio risco um comportamento mais específico de renegociação de dívida – há ainda uma oportunidade nessas contas de identificar mais variáveis que possam melhor definir seu comportamento e acrescentá-las no modelo de modo a melhorar a separação obtida.

Para cada ação tomada com um cliente inadimplente, é imprescindível seu registro no sistema que alimenta o banco de dados. Esta informação poderá ser utilizada, posteriormente, para manutenção e atualização do modelo, de forma que ele incorpore as mudanças que acontecem no ambiente de negócios, em particular na inadimplência.

Algumas das possíveis variáveis que poderiam pertencer ao modelo seriam estes resultados das próprias ações de cobrança, como por exemplo: se o cliente foi localizado com facilidade, quantas vezes foi acionado antes de fazer um pagamento ou negociar um acordo, se foi concedido desconto, etc. Porém, como já mencionado aqui, no momento da construção do modelo, não foi possível a inclusão destas variáveis específicas de cobrança no estudo, pois não dispúnhamos das mesmas devido à migração de sistema pela qual passou a instituição aqui estudada.

Ao longo da implementação é possível, também, estabelecer estratégias diferenciadas para cada parte da carteira de inadimplentes. Trata-se de dividir a mesma população em que se aplica o modelo de *charge off scoring* em duas partes: a primeira utilizará a estratégia já desenhada pela empresa para cada categoria de risco (faixa de score) e que ela acredita ser a melhor, chamada de estratégia campeã (momentaneamente as estratégias da Tabela 5.4). Para a outra parte pequena da mesma carteira (10 a 20%, aproximadamente) estabelece-se outra estratégia com ações diferentes para cada faixa de score. Ela é chamada de estratégia desafiadora.

Durante um certo período, observar-se-á a implementação das duas estratégias e comparar-se-ão as duas metodologias de tratamento deste grupo específico dos clientes inadimplentes (com 91 a 120 dias em atraso). Se a estratégia desafiadora apresentar performance superior à campeã, a primeira passa a ser a nova estratégia campeã. Pode-

se, a partir daí, desenvolver uma nova desafiadora e novamente aplicá-la a outra parte da mesma carteira. O ciclo de implementação e análise então se repete.

5.6 Manutenção

O gerenciamento das informações do modelo de *charge off scoring* é extremamente crítico para o sucesso da sua implementação. A utilização de relatórios gerenciais permite verificar se os objetivos da implementação do modelo estão sendo atendidos e também se a performance das estratégias provenientes dele está acima ou abaixo das expectativas.

As principais informações a serem gerenciadas referem-se ao resultado obtido por cada faixa de pontuação, comparando-se com as previsões estabelecidas, através da análise das perdas e recuperação para cada categoria de risco. Além disso, permite verificar qual das estratégias, campeão ou desafiadora, está gerando maior resultado para a empresa.

O objetivo, portanto, deste sistema de gerenciamento de informações não é somente avaliar o modelo, mas também as estratégias alternativas aplicadas a cada categoria de risco criada por ele.

6 CONCLUSÃO

Os objetivos iniciais de construção de um modelo para recuperação de créditos vencidos, antes que estes sejam contabilizados como prejuízo pela instituição, puderam ser plenamente atingidos através da aplicação das técnicas estatísticas aqui apresentadas. Isto pode ser diretamente medido pelo índice de separação obtido tanto para a base de dados de desenvolvimento (35%), quanto para a base de validação (33,5%). Outra questão que pode comprovar isto são os resultados obtidos com as novas estratégias que puderam ser implementadas, com a confecção do modelo (exposto no capítulo 5).

Estes resultado se resumem à seguinte equação:

$$\pi(x) = \frac{e^{0,1367 + 0,0009.SAL DO_T - 1 - 0,0014.SAL DO_T + 0,0001.LIM_T - 1,9939.UTL IM_T - 1 + 1,7587.UTL IM_T + 0,0289.MUL TA_T - 1 - 0,6315.beh_T4 - 0,7320.gr_1 - 0,5315.gr_2 - 0,3551.gr_3 - 0,2998.gr_4 + 0,2628.gr_10 + 0,4396.gr_11}}{1 + e^{0,1367 + 0,0009.SAL DO_T - 1 - 0,0014.SAL DO_T + 0,0001.LIM_T - 1,9939.UTL IM_T - 1 + 1,7587.UTL IM_T + 0,0289.MUL TA_T - 1 - 0,6315.beh_T4 - 0,7320.gr_1 - 0,5315.gr_2 - 0,3551.gr_3 - 0,2998.gr_4 + 0,2628.gr_10 + 0,4396.gr_11}}$$

Podemos identificar que , de acordo com o que foi descrito na Seção 5.1, os coeficientes indicam que uma conta terá maiores chances de não ter seu saldo contabilizado como prejuízo quando apresenta:

- existência de multa em T-1;
- maiores valores de limite de crédito concedido;
- maior utilização do limite de crédito em T;
- menor utilização do limite de crédito em T-1;
- menor saldo devido em T;
- maior saldo devido em T-1;
- menores riscos de inadimplir e;
- reside em locais que sejam os das regiões 10 e 11.

Com isso, verifica-se que o processo de modelagem proporcionou resultados financeiramente realistas para todas as variáveis independentes que o compõem (esta análise foi detalhada no quinto capítulo).

6.1 Próximos Passos

Um processo pode ser definido como uma combinação de pessoas, matéria-prima, máquinas, informações e procedimentos que combinados produzem um resultado que pode ser um produto ou serviço. Assim, podemos perceber que a atividade de recuperação de créditos vencidos encaixa-se perfeitamente na definição anterior, de modo que o modelo construído pode ser encarado como um ciclo de melhoria para o processo, de acordo com a Figura 6.1:

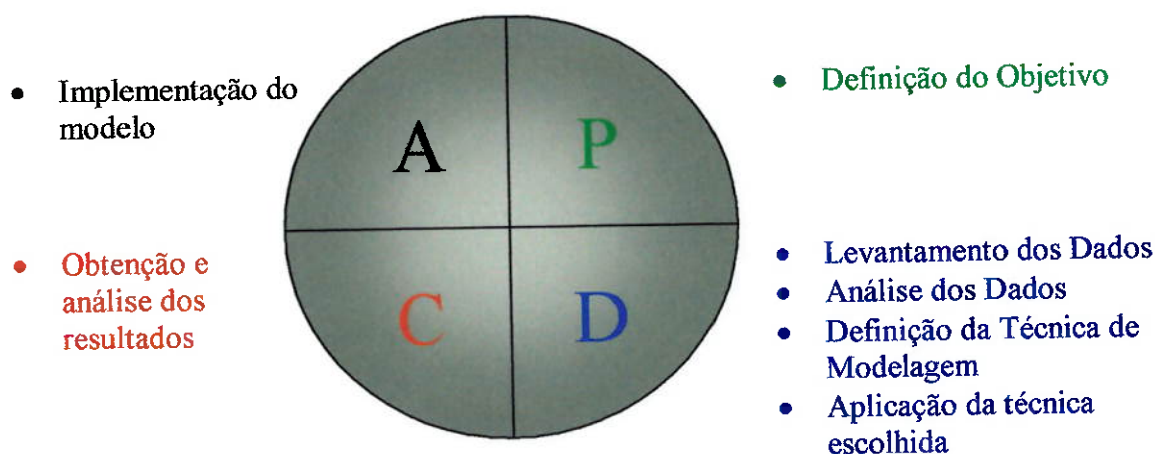


Figura 6.1 – Ciclo de Melhoria de Processo: Construção do Modelo

Diante disso, uma vez concebido este modelo, passa a ser uma atividade de suma importância sua constante atualização, de modo que seja implementado um novo ciclo de melhoria, desta vez contínuo, o qual pode ser verificado na Figura 6.2:

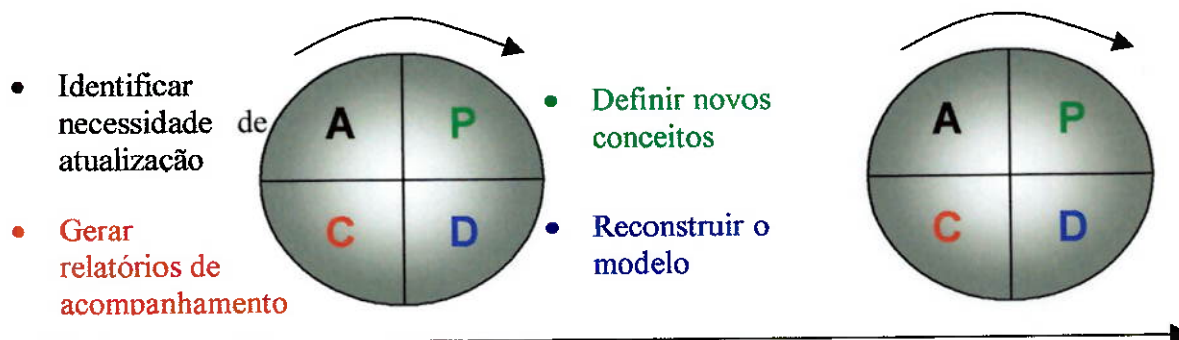


Figura 6.2 – Ciclo contínuo de melhoria: Atualização do modelo

Este ciclo deve ser “rodado” constantemente, de maneira a evitar que o modelo torne-se desatualizado frente às freqüentes mudanças da economia e do perfil do inadimplente, deixando, por conseqüência, de cumprir sua função de predição.

Para tanto, o acompanhamento do índice de separação do modelo (K-S) se faz fundamental, requerendo, quando necessário, que novas informações (variáveis) possam ser adicionadas e, eventualmente, algumas possam ser desconsideradas.

Assim, utilizando-se um programa desenvolvido em SAS poderemos extrair alguns relatórios mensais de acompanhamento do modelo, os quais conterão a distribuição das contas “boas” e “más” por classificação criada (altíssimo a baixíssimo risco de prejuízo), o índice de separação destas contas escoradas (K-S), bem como os resultados financeiros das contas classificadas (alguns destes foram expostos nas Seções 5.3, 5.4 e 5.5).

6.2 Considerações Finais

A Estatística é um conjunto de conceitos e técnicas utilizadas no planejamento da coleta de dados e sua análise, buscando a solução de questões reais. Com uma freqüência progressivamente maior, as empresas têm baseado suas decisões em análises estatísticas, o que faz com que os profissionais tenham de adquirir algum conhecimento

dessas técnicas. Sua aplicação em situações ou processos em que a incerteza e a variabilidade estão presentes torna essa disciplina de grande utilidade.

Porém, de nada adianta confeccionarmos o melhor modelo do ponto de vista estatístico se o mesmo não puder ser disponibilizado como ferramenta àqueles que precisam utilizá-lo; talvez esta seja a principal lição extraída deste projeto. O trabalho do engenheiro não se resume a encontrar uma solução para um problema real, mas sim permitir que esta possa ser perfeitamente aplicada à questão, fazendo com que os resultados esperados possam efetivamente ser alcançados.

Assim, os resultados deste trabalho ultrapassam os limites do que foi apresentado no capítulo anterior. As habilidades de atuação em processos de identificação da necessidade, busca de informações, análise, tomada de decisão e finalmente, implementação desenvolvidas através da construção do modelo são características importantíssimas atualmente para qualquer engenheiro; tudo isto, nos dá a certeza de que todos os conhecimentos adquiridos com este estudo não se restringem ao tema da recuperação de crédito, estes envolvem todas as oportunidades e desafios que irão permear a vida profissional de um Engenheiro de Produção.

7 BIBLIOGRAFIA

ASSOCIAÇÃO BRASILEIRA DAS EMPRESAS DE CARTÕES DE CRÉDITO E SERVIÇOS (ABECS). São Paulo: Disponível em: <<http://www.abecs.org.br>>. Acesso em: 10 de abr. 2001.

ASSOCIAÇÃO DAS EMPRESAS DE RECUPERAÇÃO DE CRÉDITO (ASERC). **Pesquisa sobre motivos da inadimplência**. São Paulo: ASERC, 2000.

BANCO CENTRAL DO BRASIL. Brasília. 2001. **Site na Internet do Banco Central do Brasil**. Disponível em: <www.bcb.gov.br>. Acesso em: 04 de abr. 2001.

CAOUILLE, J.B., ALTMAN, E.I., NARAYANAN, P. **Gestão do Risco de Crédito: o próximo grande desafio financeiro**; tradução de Alan Hastings; revisão técnica João Carlos Douat. Rio de Janeiro: Qualitymark Ed., 1999.

CORDELL, L. et al. Early Indicator: Design and Validation of a Comprehensive Behavior Scoring System for Delinquent Mortgages. In: CORDELL, L. et al. **Credit Risk Modeling: Design and Application**. Elizabeth Mays Editor, 1998.

COSTA NETO, Pedro Luís de Oliveira. **Estatística**. São Paulo: Edgard Blücher, 1977.

DOUAT, J.C. e SICSÚ, A.L. Reflexões sobre o desenvolvimento de sistema de classificação do risco de crédito no Brasil – Um estudo de Caso. In: 2º Fórum Internacional de Crédito SERASA, São Paulo, 1999. **Anais**. São Paulo: SERASA, 1999.

GRUPO UNIDOS. São Paulo. 2001. **Apresenta os serviços da empresa, além de pesquisas sobre inadimplência**. Disponível em: <<http://www.grupounidos.com.br>>. Acesso em: 13 de jul. 2001.

HAIR, Joseph F. **Multivariate Data Analysis with readings**. Prentice Hall: Englewoodcliffs, N.J., 1995.

HOSMER JR, David W., LEMESHOW, Stanley. **Applied Logistic Regression**. Wiley-Interscience: Massachussets, USA, 1989.

MCCA HILL, L.J. Organizational Issues in Building and Maintaining Credit Risk Models. In: CORDELL, L. et al. **Credit Risk Modeling: Design and Application**. Elizabeth Mays Editor, 1998.

MAKUCH, W.M. The Basics of a Better Application Score. In: CORDELL, L. et al. **Credit Risk Modeling: Design and Application**. Elizabeth Mays Editor, 1998.

MALAGA, T., PRANDINI, J.C. e RABBAT, M. Default: Applying the Economic Theory of Information. In: 2º Fórum Internacional de Crédito SERASA, São Paulo, 1999. **Anais**. São Paulo: SERASA, 1999.

MOORE, D.S. The basic practice of statistics. New York: W. H. Freeman and Company, 1994.

NAKAMURA, Daniel. **Aplicações de Técnicas Estatísticas na Construção de um Modelo de Risco de Crédito**. Trabalho de Formatura. São Paulo: EPUSP – Departamento de Engenharia de Produção, 2000.

PANORAMA SETORIAL E GAZETA MERCANTIL. **Análise Setorial – Cartões de Crédito e Débito**. São Paulo, 1999.

SERASA. São Paulo. 2001. **Apresenta os serviços da empresa, além de pesquisas sobre o mercado de crédito**. Disponível em: <www.serasa.com.br>. Acesso em: 10 de abr. 2001.

SILVA, José Pereira da. Gestão e análise de risco de crédito. São Paulo: Atlas, 1997.

SOUZA, Rodnei Bernardino. O Modelo de Collection Scoring como Ferramenta para a Gestão Estratégica do Risco de Crédito. Dissertação de MBA apresentada ao Curso de Pós Graduação. São Paulo: EAESP/FGV – Área de Concentração: Risco de Crédito, 2000.

TEC COBRA – ESCRITÓRIO DE COBRANÇA. **Pesquisa de inadimplência de agosto de 2001.** São Paulo: Tec Cobra, 2001.

8 ANEXOS

8.1 Anexo I - Resolução 2682

Retirada do site <http://www.bcb.gov.br>.

RESOLUCAO N. 002682

Dispoe sobre criterios de classificacao das operacoes de credito e regras para constituicao de provisao para creditos de liquidacao duvidosa.

O BANCO CENTRAL DO BRASIL, na forma do art. 9. da Lei n. 4.595, de 31 de dezembro de 1964, torna publico que o CONSELHO MONETARIO NACIONAL, em sessao realizada em 21 de dezembro de 1999, com base no art. 4., incisos XI e XII, da citada Lei,

R E S O L V E U:

Art. 1. Determinar que as instituicoes financeiras e demais instituicoes autorizadas a funcionar pelo Banco Central do Brasil devem classificar as operacoes de credito, em ordem crescente de risco, nos seguintes niveis:

- I - nivel AA;
- II - nivel A;
- III - nivel B;
- IV - nivel C;
- V - nivel D;
- VI - nivel E;
- VII - nivel F;
- VIII - nivel G;
- IX - nivel H.

Art. 2. A classificacao da operacao no nivel de risco correspondente e de responsabilidade da instituicao detentora do credito e deve ser efetuada com base em criterios consistentes e verificaveis, amparada por informacoes internas e externas, contemplando, pelo menos, os seguintes aspectos:

I - em relação ao devedor e seus garantidores:

- a) situação econômico-financeira;
- b) grau de endividamento;
- c) capacidade de geração de resultados;
- d) fluxo de caixa;
- e) administração e qualidade de controles;
- f) pontualidade e atrasos nos pagamentos;
- g) contingências;
- h) setor de atividade econômica;
- i) limite de crédito;

II - em relação à operação:

- a) natureza e finalidade da transação;
- b) características das garantias, particularmente quanto à suficiência e liquidez;
- c) valor.

Parágrafo único. A classificação das operações de crédito de titularidade de pessoas físicas deve levar em conta, também, as situações de renda e de patrimônio bem como outras informações cadastrais do devedor.

Art. 3. A classificação das operações de crédito de um mesmo cliente ou grupo econômico deve ser definida considerando aquela que apresentar maior risco, admitindo-se excepcionalmente classificação diversa para determinada operação, observado o disposto no art. 2., inciso II.

Art. 4. A classificação da operação nos níveis de risco de que trata o art. 1. deve ser revista, no mínimo:

I - mensalmente, por ocasião dos balancetes e balanços, em função de atraso verificado no pagamento de parcela de principal ou de encargos, devendo ser observado o que segue:

- 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;

II - com base nos critérios estabelecidos nos arts. 2. e 3.:

a) a cada seis meses, para operações de um mesmo cliente ou grupo econômico cujo montante seja superior a 5% (cinco por cento) do patrimônio líquido ajustado;

b) uma vez a cada doze meses, em todas as situações, exceto na hipótese prevista no art. 5..

Parágrafo 1. As operações de adiantamento sobre contratos de câmbio, as de financiamento a importação e aquelas com prazos inferiores a um mês, que apresentem atrasos superiores a trinta dias, bem como o adiantamento a depositante a partir de trinta dias de sua ocorrência, devem ser classificadas, no mínimo, como de risco nível G.

Parágrafo 2. Para as operações com prazo a decorrer superior a 36 meses admite-se a contagem em dobro dos prazos previstos no inciso I.

Parágrafo 3. O não atendimento ao disposto neste artigo implica a reclassificação das operações do devedor para o risco nível H, independentemente de outras medidas de natureza administrativa.

Art. 5. As operações de crédito contratadas com cliente cuja responsabilidade total seja de valor inferior a R\$ 50.000,00 (cinquenta mil reais) podem ter sua classificação revista de forma automática unicamente em função dos atrasos consignados no art. 4., inciso I, desta Resolução, observado que deve ser mantida a classificação original quando a revisão corresponder a nível de menor risco.

Parágrafo 1. O Banco Central do Brasil poderá alterar o valor de que trata este artigo.

Parágrafo 2. O disposto neste artigo aplica-se às operações contratadas até 29 de fevereiro de 2000, observados o valor referido no caput e a classificação, no mínimo, como de risco nível A.

Art. 6. A provisão para fazer face aos créditos de liquidação duvidosa deve ser constituída mensalmente, não podendo ser inferior ao somatório decorrente da aplicação dos percentuais a seguir mencionados, sem prejuízo da responsabilidade dos administradores das instituições pela constituição de provisão em montantes suficientes para fazer face a perdas prováveis na realização dos créditos:

- I - 0,5% (meio por cento) sobre o valor das operações classificadas como de risco nível A;

II - 1% (um por cento) sobre o valor das operacoes classificadas como de risco nivel B;

III - 3% (tres por cento) sobre o valor das operacoes classificadas como de risco nivel C;

IV - 10% (dez por cento) sobre o valor das operacoes classificados como de risco nivel D;

V - 30% (trinta por cento) sobre o valor das operacoes classificados como de risco nivel E;

VI - 50% (cinquenta por cento) sobre o valor das operacoes classificados como de risco nivel F;

VII - 70% (setenta por cento) sobre o valor das operacoes classificados como de risco nivel G;

VIII - 100% (cem por cento) sobre o valor das operacoes classificadas como de risco nivel H.

Art. 7. A operacao classificada como de risco nivel H deve ser transferida para conta de compensacao, com o correspondente debito em provisao, apos decorridos seis meses da sua classificacao nesse nivel de risco, nao sendo admitido o registro em periodo inferior.

Paragrafo unico. A operacao classificada na forma do disposto no caput deste artigo deve permanecer registrada em conta de compensacao pelo prazo minimo de cinco anos e enquanto nao esgotados todos os procedimentos para cobranca.

Art. 8. A operacao objeto de renegociacao deve ser mantida, no minimo, no mesmo nivel de risco em que estiver classificada, observado que aquela registrada como prejuizo deve ser classificada como de risco nivel H.

Paragrafo 1. Admite-se a reclassificacao para categoria de menor risco quando houver amortizacao significativa da operacao ou quando fatos novos relevantes justificarem a mudanca do nivel de risco.

Paragrafo 2. O ganho eventualmente auferido por ocasio da renegociacao deve ser apropriado ao resultado quando do seu efetivo recebimento.

Paragrafo 3. Considera-se renegociacao a composicao de divida, a prorrogacao, a novacao, a concessao de nova operacao para liquidacao parcial ou integral de operacao anterior ou qualquer outro tipo de acordo que implique na alteracao nos prazos de vencimento ou nas condicoes de pagamento originalmente pactuadas.

Art. 9. E vedado o reconhecimento no resultado do periodo de receitas e encargos de qualquer natureza relativos a operacoes de credito que apresentem atraso igual ou superior a sessenta dias, no pagamento de parcela de principal ou encargos.

Art. 10. As instituições devem manter adequadamente documentadas sua política e procedimentos para concessão e classificação de operações de crédito, os quais devem ficar a disposição do Banco Central do Brasil e do auditor independente.

Parágrafo único. A documentação de que trata o caput deste artigo deve evidenciar, pelo menos, o tipo e os níveis de risco que se dispõe a administrar, os requerimentos mínimos exigidos para a concessão de empréstimos e o processo de autorização.

Art. 11. Devem ser divulgadas em nota explicativa as demonstrações financeiras informações detalhadas sobre a composição da carteira de operações de crédito, observado, no mínimo:

I - distribuição das operações, segregadas por tipo de cliente e atividade econômica;

II - distribuição por faixa de vencimento;

III - montantes de operações renegociadas, lançados contra prejuízo e de operações recuperadas, no exercício.

Art. 12. O auditor independente deve elaborar relatório circunstanciado de revisão dos critérios adotados pela instituição quanto à classificação nos níveis de risco e de avaliação do provisionamento registrado nas demonstrações financeiras.

Art. 13. O Banco Central do Brasil poderá baixar normas complementares necessárias ao cumprimento do disposto nesta Resolução, bem como determinar:

I - reclassificação de operações com base nos critérios estabelecidos nesta Resolução, nos níveis de risco de que trata o art. 1.º;

II - provisionamento adicional, em função da responsabilidade do devedor junto ao Sistema Financeiro Nacional;

III - providências saneadoras a serem adotadas pelas instituições, com vistas a assegurar a sua liquidez e adequada estrutura patrimonial, inclusive na forma de alocação de capital para operações de classificação considerada inadequada;

IV - alteração dos critérios de classificação de créditos, de contabilização e de constituição de provisões;

V - teor das informações e notas explicativas constantes das demonstrações financeiras;

VI - procedimentos e controles a serem adotados pelas instituições.

Art. 14. O disposto nesta Resolução se aplica também às ope-

rações de arrendamento mercantil e a outras operações com características de concessão de crédito.

Art. 15. As disposições desta Resolução não contemplam os aspectos fiscais, sendo de inteira responsabilidade da instituição a observância das normas pertinentes.

Art. 16. Esta Resolução entra em vigor na data da sua publicação, produzindo efeitos a partir de 1. de março de 2000, quando ficarão revogadas as Resoluções n.s 1.748, de 30 de agosto de 1990, e 1.999, de 30 de junho de 1993, os arts. 3. e 5. da Circular n. 1.872, de 27 de dezembro de 1990, a alínea "b" do inciso II do art. 4. da Circular n. 2.782, de 12 de novembro de 1997, e o Comunicado n. 2.559, de 17 de outubro de 1991.

Brasília, 21 de dezembro de 1999

Arminio Fraga Neto
Presidente

8.2 Anexo II – Valores Críticos da Distribuição Qui-quadrado

Extraída de Moore (1995).

df	p											
	0,25	0,20	0,15	0,10	0,05	0,025	0,02	0,01	0,005	0,0025	0,001	0,0005
1	1,32	1,64	2,07	2,71	3,84	5,02	5,41	6,63	7,88	9,14	10,83	12,12
2	2,77	3,22	3,79	4,61	5,99	7,38	7,82	9,21	10,60	11,98	13,82	15,20
3	4,11	4,64	5,32	6,25	7,81	9,35	9,84	11,34	12,84	14,32	16,27	17,73
4	5,39	5,99	6,74	7,78	9,49	11,14	11,67	13,28	14,86	16,42	18,47	20,00
5	6,63	7,29	8,12	9,24	11,07	12,83	13,39	15,09	16,75	18,39	20,51	22,11
6	7,84	8,56	9,45	10,64	12,59	14,45	15,03	16,81	18,55	20,25	22,46	24,10
7	9,04	9,80	10,75	12,02	14,07	16,01	16,62	18,48	20,28	22,04	24,32	26,02
8	10,22	11,03	12,03	13,36	15,51	17,53	18,17	20,09	21,95	23,77	26,12	27,87
9	11,39	12,24	13,29	14,66	16,92	19,02	19,68	21,67	23,59	25,46	27,88	29,67
10	12,55	13,44	14,53	15,99	18,31	20,48	21,16	23,21	25,19	27,11	29,59	31,42
11	13,70	14,63	15,77	17,28	19,68	21,92	22,62	24,72	26,76	28,73	31,26	33,14
12	14,85	15,81	16,99	18,55	21,03	23,34	24,05	26,22	28,30	30,32	32,91	34,82
13	15,98	16,98	18,20	19,81	22,36	24,74	25,47	27,69	29,82	31,88	34,53	36,48
14	17,12	18,15	19,41	21,06	23,68	26,12	26,87	29,14	31,32	33,43	36,12	38,11
15	18,25	19,31	20,60	22,31	25,00	27,49	28,26	30,58	32,80	34,95	37,70	39,72
16	19,37	20,47	21,79	23,54	26,30	28,85	29,63	32,00	34,27	36,46	39,25	41,31
17	20,49	21,61	22,98	24,77	27,59	30,19	31,00	33,41	35,72	37,95	40,79	42,88
18	21,60	22,76	24,16	25,99	28,87	31,53	32,35	34,81	37,16	39,42	42,31	44,43
19	22,72	23,90	25,33	27,20	30,14	32,85	33,69	36,19	38,58	40,88	43,82	45,97
20	23,83	25,04	26,50	28,41	31,41	34,17	35,02	37,57	40,00	42,34	45,31	47,50
21	24,93	26,17	27,66	29,62	32,67	35,48	36,34	38,93	41,40	43,78	46,80	49,01
22	26,04	27,30	28,82	30,81	33,92	36,78	37,66	40,29	42,80	45,20	48,27	50,51
23	27,14	28,43	29,98	32,01	35,17	38,08	38,97	41,64	44,18	46,62	49,73	52,00
24	28,24	29,55	31,13	33,20	36,42	39,36	40,27	42,98	45,56	48,03	51,18	53,48
25	29,34	30,68	32,28	34,38	37,65	40,65	41,57	44,31	46,93	49,44	52,62	54,95
26	30,43	31,79	33,43	35,56	38,89	41,92	42,86	45,64	48,29	50,83	54,05	56,41
27	31,53	32,91	34,57	36,74	40,11	43,19	44,14	46,96	49,64	52,22	55,48	57,86
28	32,62	34,03	35,71	37,92	41,34	44,46	45,42	48,28	50,99	53,59	56,89	59,30
29	33,71	35,14	36,85	39,09	42,56	45,72	46,69	49,59	52,34	54,97	58,30	60,73
30	34,80	36,25	37,99	40,26	43,77	46,98	47,96	50,89	53,67	56,33	59,70	62,16
40	45,62	47,27	49,24	51,81	55,76	59,34	60,44	63,69	66,77	69,70	73,40	76,09
50	56,33	58,16	60,35	63,17	67,50	71,42	72,61	76,15	79,49	82,66	86,66	89,56
60	66,98	68,97	71,34	74,40	79,08	83,30	84,58	88,38	91,95	95,34	99,61	102,70
80	88,13	90,41	93,11	96,58	101,90	106,60	108,10	112,30	116,30	120,10	124,80	128,30
100	109,10	111,70	114,70	118,50	124,30	129,60	131,10	135,80	140,20	144,30	149,40	153,20