**ANÁLISE DA INADIMPLÊNCIA EM FINANCIAMENTOS DE
LONGO PRAZO NO ESPÍRITO SANTO: UM ESTUDO
DE CASO PARA O SETOR HOTELEIRO**

MAURÍCIO CÉZAR DUQUE

Economista

Orientador: Prof. Dr. **PAULO FERNANDO CIDADE DE ARAÚJO**

Dissertação apresentada à Escola Superior de
Agricultura "Luiz de Queiroz", Universidade de São
Paulo, para obtenção do título de Mestre em
Ciências, Área de Concentração: Economia
Aplicada.

PIRACICABA

Estado de São Paulo – Brasil

Fevereiro- 2000

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - Campus "Luiz de Queiroz"/USP

Duque, Maurício César

Análise da inadimplência em financiamentos de longo prazo no Espírito Santo: um estudo de caso para o setor hoteleiro / Maurício César Duque. - - Piracicaba, 2000.
93 p.

Dissertação (mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2000.
Bibliografia.

1. Análise econômica 2. Desenvolvimento econômico 3. Economia agrícola 4.
Estabilização econômica 5. Financiamento empresarial 6. Hotel 7. Indústria agrícola
8. Modelo matemático I. Título

CDD 338.1

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

Dedido aos meus pais, Demerval e Ireni, a minha afilhada, Vanessa,
a minha esposa, Luciana, e ao meu filho, Leonardo

AGRADECIMENTOS

A realização deste trabalho é consequência da colaboração irrestrita que obtive de amigos e familiares.

Ao Departamento de Economia da ESALQ/USP e ao CNPq, fica o meu agradecimento pela oportunidade de realizar o curso de mestrado.

Ao Banco de Desenvolvimento do Espírito Santo S/A (BANDES) e ao Banco Santos Neves S/A agradeço pela atenção ímpar e pelas informações fornecidas para a realização deste trabalho.

Aos professores Adriano, Caixeta, Bacha, João Martinez, Bento, Ana Kassouf, Paulo Cidade e Pedro Marques, meus agradecimentos pelos ensinamentos transmitidos ao longo do curso.

Um agradecimento especial ao professor Paulo Cidade pela orientação e ao professor Adriano pelas inúmeras idéias e contribuições na fase inicial deste trabalho. Aos professores Caixeta e Míriam fica também o agradecimento pelas importantes contribuições dadas na fase final do trabalho.

Aos funcionários do Departamento de Economia, em especial às funcionárias Maieli e Luciane, agradeço pela dedicação e disposição ao longo do curso.

Aos colegas Hemildo, Márcio, Rinaldo, Rogério, Daniela, Stivilane, Margareth, Lílian, Pedro e Marcelo agradeço pela amizade e pelos momentos fraternos que compartilhamos.

Um agradecimento especial aos amigos de todos os momentos, Augusto, Fábio, Felipe, Renato e Valter. Na verdade, mostraram-se muito mais que amigos,

mostraram-se irmãos. Irmãos com inteligência, caráter e bondade como poucos têm o privilégio de possuir.

Aos meus pais, Demerval e Ireni, aos meus irmãos Marcelo e Márcia, e a minha afilhada, Vanessa, agradeço com lágrimas nos olhos, pois ainda lembro da saudade sufocante e do amor transmitido durante os anos de "separação".

A minha esposa Luciana, agradeço a confiança, o companheirismo, o carinho e o amor. Sem a "minha Lú", com seu coração enorme e seu amor sincero, eu jamais encontraria o caminho para a felicidade.

Ao meu filho Leonardo, agradeço pelo sorriso e pelo olhar de quem ainda é um anjo, que não entende o que se passa, mas nos contagia com a esperança de um mundo mais justo.

E, finalmente, por toda a minha vida, agradeço a Deus.

SUMÁRIO

	Página
LISTA DE FIGURAS.....	viii
LISTA DE TABELAS.....	ix
RESUMO.....	x
SUMMARY.....	xii
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	4
2.1. A inadimplência no sistema financeiro nacional.....	4
2.2. O "estado da arte" no gerenciamento de crédito.....	12
3. REFERENCIAL TEÓRICO.....	18
3.1. O Banco de Desenvolvimento do Espírito Santo S/A - BANDES.....	19
3.1.1. A Inadimplência no BANDES.....	23
3.2. Modelos de resposta qualitativa.....	26
3.2.1. Modelo utilizando probabilidade linear.....	28
3.2.1.1. Problemas com modelos de probabilidade linear.....	29
3.2.2. Modelos lógite e próbite.....	31
3.2.2.1. Estimação dos parâmetros.....	33
3.2.3. Considerações sobre modelos de resposta qualitativa.....	35
3.3. Modelo baseado no conceito de entropia.....	36
3.4. Considerações sobre o uso de modelos para previsão de inadim- plência.....	37
4. ANÁLISE DOS RESULTADOS.....	40

4.1. Modelos de resposta qualitativa.....	41
4.1.1. Modelo de probabilidade linear.....	42
4.1.2. Modelos logite e probite.....	45
4.2. Modelo baseado no conceito de entropia.....	49
5. CONCLUSÕES.....	55
REFERÊNCIAS BIBLIOGRÁFICAS.....	60
APÊNDICE 1: Dados.....	64
APÊNDICE 2: Estimativas das regressões do modelo de probabilidade Linear.....	65
APÊNDICE 3: Estimativas das regressões do modelo probite.....	66
APÊNDICE 4: Estimativas das regressões do modelo logite.....	68
APÊNDICE 5: "Árvores Discriminatórias" do modelo baseado no conceito de entropia.....	70

LISTA DE FIGURAS

	Página
Figura 1. Empréstimos do Setor Financeiro.....	6
Figura 2. Inadimplência nos Empréstimos do Setor Financeiro.....	7
Figura 3. Inadimplência nos Empréstimos do Setor Financeiro por Atividade.....	9
Figura 4. BANDES: Investimentos Aprovados por Fonte de Financiamento.....	21
Figura 5. BANDES: Investimentos Aprovados por Objetivo.....	22
Figura 6. BANDES: Investimentos Aprovados por Atividade.....	22
Figura 7. Função de Distribuição Cumulativa.....	31
Figura 8. “Árvore Discriminatória I” para Projetos de Hotelaria Aprovados pelo BANDES.....	52

LISTA DE TABELAS

	Página
Tabela 1. Evolução da Receita Inflacionária: 1990-1997.....	5
Tabela 2. BANDES: Investimentos Aprovados: 1991/1996.....	19
Tabela 3. Níveis de Inadimplência (%): Média Nacional e BANDES.....	23
Tabela 4. BANDES: Taxa de Inadimplência por Atividade (Base: Março/1997).....	25
Tabela 5. Variáveis Explicativas do Modelo de Resposta Qualitativa.....	41
Tabela 6. Variáveis Descartadas.....	42
Tabela 7. Estimativas dos Parâmetros do Modelo de Probabilidade Linear.....	43
Tabela 8. Modelo de Probabilidade Linear: Observação x Previsão.....	44
Tabela 9. Estimativas dos Parâmetros do Modelo Próbite.....	45
Tabela 10. Estimativas dos Parâmetros do Modelo Lógite.....	46
Tabela 11. Modelo Próbite: Observação x Previsão.....	47
Tabela 12. Modelo Lógite: Observação x Previsão.....	48
Tabela 13. Variáveis Explicativas do Modelo Utilizando o Conceito de Entropia.....	50
Tabela 14. Análise Estatística Descritiva das Variáveis Explicativas.....	57
Tabela 15. Coeficientes de Regressão.....	58

ANÁLISE DA INADIMPLÊNCIA EM FINANCIAMENTOS DE LONGO PRAZO NO ESPÍRITO SANTO: UM ESTUDO DE CASO PARA O SETOR HOTELEIRO

Autor: MAURÍCIO CÉZAR DUQUE

Orientador: Prof. Dr. PAULO FERNANDO CIDADE DE ARAÚJO

RESUMO

O presente estudo tem como finalidade o desenvolvimento de um modelo para identificação e quantificação das variáveis econômicas que melhor explicam o fenômeno da inadimplência em financiamentos de longo prazo concedidos pelo Banco de Desenvolvimento do Espírito Santo S/A [BANDES] ao segmento de hotelaria.

A importância de modelos confiáveis para a indentificação *ex-ante* da probabilidade de ocorrência de inadimplência em créditos de longo prazo acentuou-se ainda mais no período pós Plano Real, quando o volume de empréstimos do sistema financeiro nacional ao setor privado cresceu a taxas significativas, substituindo, temporariamente, a chamada receita inflacionária.

Os modelos sugeridos foram construídos com base nos conceitos de entropia e resposta qualitativa, com destaque para os modelos de probabilidade linear, próbite e lógite. Sendo que enquanto o modelo baseado no conceito de entropia permitiu apenas uma ordenação das variáveis escolhidas, os modelos de resposta qualitativa permitiram a quantificação das mesmas, tendo sido, portanto, mais eficientes.

O ganho de informação obtido com a utilização de tais modelos foi significativo. No caso do modelo de probabilidade linear, o percentual de classificação correta ficou em 80%, enquanto nos modelos próbite e lógite este percentual alcançou 82%. Devendo-se ressaltar que a significância estatística do modelo e dos parâmetros indica que as variáveis selecionadas são importantes para explicar o fenômeno da inadimplência na carteira de empréstimos do BANDES ao segmento de hotelaria.

As variáveis escolhidas para a explicação do fenômeno foram: investimento total [valores em R\$]; capital de terceiros [valores em R\$]; taxa de juros média anual nos financiamentos [valores em %]; prazo de amortização do financiamento [valores em meses]; faturamento médio anual previsto [valores em R\$]; e investimento por unidade habitacional [valores em R\$].

Finalmente, cabe ressaltar a sugestão pela utilização do modelo Lógite para problemas como o apresentado neste trabalho, em virtude, principalmente, de sua eficiência e facilidade operacional.

ANALYSIS OF DEFAULT IN LOANS OF LONG TERM IN ESPÍRITO SANTO: A STUDY OF CASE FOR HOTEL SEGMENT

Author: MAURÍCIO CÉZAR DUQUE

Adviser: Prof. Dr. PAULO FERNANDO CIDADE DE ARAÚJO

SUMMARY

The present study has as purpose the development of a model for identification and quantification of the economic variables that best explains the phenomenon of default in loans of long term granted by Bank of Development of Espírito Santo [BANDES] to the hotel segment.

The importance of reliable models to identify “ex-ante” the probability of default occurrence in credits of long term was accentuated even more in the period after “Plano Real”, when the volume of loans of national financial system to private segment grew at a significant rate, substituting, temporarily, the named inflationary revenue.

The suggested models were built based on entropy concepts and binary choice, with prominence for linear regression, Probit and Logit models. While the model based on entropy concept just allowed an ordering of chosen variables, the models of binary choice allowed their quantification, becoming, therefore, more efficient.

The gain of information obtained using such models was significant. In case of linear regression model, the percentile of correct classification was 80%, while in Probit and Logit models this percentile reached 82%. It should be stood out that the statistical significance of both model and parameters indicates that the selected variables are important to explain the phenomenon of default in the wallet of loans of BANDES to the hostelry segment.

The variables chosen to explain the phenomenon were: total investment [values in R\$]; capital of third [values in R\$]; annual medium interest rate in financings [values in %]; period of amortization of financing [values in months]; foreseen annual medium revenue [values in R\$]; and investment per habitational unit [values in R\$].

Finally, it fits to emphasize the suggestion for using the Logit model in problems as presented in this work, due to, mainly, its efficiency and its operational easiness.

1. INTRODUÇÃO

O problema do crescimento dos níveis de inadimplência vem afetando diversos segmentos da economia brasileira. O sistema financeiro nacional, que passa por um momento de forte alteração em sua estrutura, é um dos setores onde tal problema se manifesta de forma mais evidente e preocupante. Desde a adoção do Plano Real, em junho de 1994, os níveis de inadimplência junto ao sistema financeiro cresceram a taxas superiores a 100%.

Nesse período verificou-se uma significativa expansão nos saldos das operações de crédito das instituições financeiras. Entretanto, face as características do processo de concessão de crédito do período anterior, influenciadas sobretudo pelas elevadas taxas de inflação, percebeu-se que os critérios utilizados por essas instituições para a classificação *ex-ante* do risco de emprestar eram extremamente falhos.

A necessidade de se conhecer as variáveis micro e macroeconômicas que afetam as condições de pagamento de uma empresa junto aos seus credores preocupa sobretudo as instituições financeiras. Os critérios para aprovação de crédito estão sendo reformulados, visando a redução da inadimplência e adequação dos "spreads" adotados, ou seja, em operações com riscos maiores utilizar-se-ia taxas de juros mais elevadas.

No caso específico do Banco de Desenvolvimento do Espírito Santo (BANDES), instituição financeira a ser analisada neste trabalho, o problema da inadimplência assume contornos ainda mais preocupantes, pois trata-se do agente responsável pela execução das políticas de incentivos fiscais e creditícios do Estado do Espírito Santo e que trabalha, quase exclusivamente, com financiamentos de longo prazo. Em dados preliminares, observa-se que os níveis de inadimplência nessa instituição encontram-se, no geral, bem superiores à média nacional e contribuem de forma significativa para os resultados negativos apresentados pelo banco nos últimos anos.

No intuito de oferecer uma contribuição para um melhor entendimento do problema, objetiva-se nesse estudo a elaboração de um modelo estatístico que possa explicar a ocorrência de créditos inadimplentes, levando-se em conta variáveis que possam contribuir direta ou indiretamente para o desenvolvimento desse fenômeno. Espera-se que esse modelo, construído a partir de técnicas relacionadas à análise de respostas qualitativas, possa ser útil para orientar o processo de decisão do BANDES.

Como as variáveis econômicas têm efeitos distintos sobre os diferentes segmentos com os quais o BANDES opera, decidiu-se então pela segmentação da análise, ou seja, a construção de modelos específicos. Neste trabalho, o modelo proposto recai sobre o segmento de hotelaria, cujo histórico de atrasos e inadimplências junto a essa instituição é significativo.

Para facilitar a análise do problema da inadimplência junto aos financiamentos concedidos pelo BANDES para o segmento de hotelaria, decidiu-se segmentar esse trabalho em cinco capítulos. No capítulo seguinte mostra-se o quadro geral da inadimplência junto ao sistema financeiro nacional, destacando os índices setoriais e a importância do problema, e, também, comenta-se alguns estudos com problema e/ou metodologia semelhantes às

que serão utilizadas neste trabalho, devendo, desde já, ressaltar a carência de estudos sobre a inadimplência em financiamentos de longo prazo voltados para pessoas jurídicas.

A seguir, no capítulo 3, faz-se uma breve apresentação sobre o BANDES, ressaltando alguns aspectos operacionais e financeiros do mesmo, procurando-se evidenciar a intensidade dos problemas que as operações inadimplentes vêm ocasionando à sua carteira de crédito. Também nesse capítulo serão apresentados os modelos de resposta qualitativa - com destaque para os modelos de probabilidade linear, modelo lógite e modelo próbite – e o modelo baseado no conceito de entropia.

No capítulo 4, apresentam-se os resultados obtidos com os modelos propostos, seus níveis de acerto e as comparações entre os mesmos. Tal capítulo será o referencial para o capítulo 5, no qual são apresentadas as considerações finais.

2. REVISÃO DE LITERATURA

2.1. A INADIMPLÊNCIA¹ NO SISTEMA FINANCEIRO NACIONAL

Com o processo de estabilização da economia, em curso desde a implantação do Plano Real, em julho de 1994, o sistema financeiro brasileiro vem atravessando momentos de intensa alteração em sua estrutura, visando, principalmente, adequar-se a novas formas de financiamento que permitam a manutenção de suas atividades.

Os seguidos anos de inflação em níveis elevados e os constantes desequilíbrios macroeconômicos incentivaram o aparecimento de inúmeras instituições financeiras que, objetivando ganhos “fáceis” com receitas inflacionárias (*float*),² utilizavam como forma de captação de recursos (depósitos e aplicações) a abertura indiscriminada de novas agências.³

¹ O Banco Central do Brasil define inadimplência como o percentual referente a créditos em atraso e em liquidação sobre o saldo total de operações de crédito. Em geral, consideram-se créditos em atraso, as operações vencidas há mais de 60 dias e reclassificadas pelo valor atualizado. Os créditos em liquidação são as operações vencidas há mais de 180 dias com garantias consideradas insuficientes e há mais de 360 dias com garantias suficientes (Resolução N° 1.748, de 30/08/1990).

² Esse ganho era propiciado: (i) pela perda do valor real dos depósitos a vista e/ou (ii) pela correção dos depósitos bancários em valores abaixo da inflação.

³ Esperava-se que as instituições financeiras comessem seu processo de ajuste pelo fechamento das agências economicamente inviáveis. No entanto, na prática tal fato não se concretizou, visto a pequena redução no número de agências entre dezembro de 1994 e dezembro de 1995 (de 17.939 para 17.865), segundo dados do Banco Central.

Obrigado(a), “float”, conforme visualiza-se na tabela abaixo, era a principal fonte de receita das instituições financeiras brasileiras. Segundo Mendonça de Barros e Almeida (1997) a estabilização de preços possibilitou uma diminuição significativa da transferência de renda dos setores não bancários para o setor bancário. Tomando-se como base a média de 1990 a 1993 (a valores de 1994) estima-se algo próximo a R\$ 19 bilhões de perda para os bancos.

Tabela 1. Evolução da Receita Inflacionária: 1990-1997

Período	Receita Inflacionária / PIB (%)	Receita Inflacionária / Valor da Produção das Instituições Bancárias (%)
1990	4,00	35,7
1991	3,90	41,3
1992	4,00	41,9
1993	4,20	35,3
1994	2,00	20,4
1995	0,50	...
1996	0,25	...
1997	0,15	...

Fontes:

1990 - 94: ANDIMA/IBGE: Sistema Financeiro - Uma análise a partir das Contas Nacionais 1990-1995.

1995 - 97: IPEA

(...) - Informações não disponíveis

A facilidade existente antes do Plano Real, qual seja, se “emprestar” dinheiro para o setor público com taxa de risco próxima a zero, foi substituída pela necessidade (e pelo risco) de se conceder crédito ao setor privado. A expansão desse tipo de crédito foi significativa, passando de um saldo de operações de cerca de R\$ 144 bilhões, em julho de 1994, para algo próximo a R\$ 276 bilhões, em janeiro de 1998, ou seja, acréscimo de aproximadamente 92% (figura 1).

Esse significativo crescimento dos empréstimos das instituições financeiras ao setor privado pôde compensar, ao menos parcial e temporariamente, a perda das receitas obtidas através da “*transferência inflacionária*”. Outra solução adotada para compensar a perda de receita oriunda do *float* foi a cobrança de tarifas pelos serviços prestados aos clientes (extratos, talões extras, etc.).

Figura 1. Empréstimos do Setor Financeiro

Fonte: Boletim do Banco Central do Brasil - Abril/1998.

Contudo, ainda segundo Mendonça de Barros e Almeida (1997), um grande problema em períodos de expansão rápida dos créditos é o aumento da vulnerabilidade das instituições financeiras. Os autores baseiam-se no fato de que esses períodos são momentos de expansão macroeconômica, quando os devedores estão transitoriamente com folga de liquidez, dificultando, assim, uma análise de risco mais rigorosa por parte dos bancos. Ademais, nesses períodos, os bancos são levados a aumentarem a carteira de crédito através da incorporação de novos clientes. Este problema é ainda maior, no período inicial

da estabilização, quando os balanços (e/ou balancetes) do período antigo são pouco informativos e muitos dos cadastros dos clientes têm de ser refeitos.

Desta forma, completam Barros e Almeida (1996), o resultado desse processo é um crescimento dos empréstimos de liquidação duvidosa e uma maior vulnerabilidade do sistema bancário, principalmente quando ocorre algum choque macroeconômico adverso. A diminuição no ritmo de crescimento da economia brasileira, verificada no segundo trimestre de 1995, confirmou essa hipótese, com o aumento substancial nos créditos inadimplentes no sistema financeiro.

A magnitude desse crescimento pode ser visualizada na figura 2, na qual percebe-se uma elevação na taxa de inadimplência de 7,58% em julho de 1994 para 15,21% em janeiro de 1998 (alcançando o máximo de 17,02% em outubro de 1997), o que representa créditos inadimplentes da ordem de R\$ 45,94 bilhões. Pode-se observar, então, que a variação da taxa de inadimplência geral no período foi superior a 100%.

Figura 2. Inadimplência nos Empréstimos do Setor Financeiro

Fonte: Boletim do Banco Central do Brasil - Abril/1998.

Obs.: Não inclui créditos em liquidação de duas instituições financeiras.

Na figura 3 apresenta-se a inadimplência por atividade econômica, segundo a classificação do BACEN, com destaque negativo nesse novo contexto econômico para a atividade “comércio”, a qual apresentou, no mesmo período, uma variação superior a 350%, passando de 6,24% para 28,22%, o que representa algo próximo a R\$ 8,7 bilhões. Sendo interessante considerar que essa taxa já esteve em 33,38%.

As demais atividades também apresentaram taxas crescentes de inadimplência. Nas atividades “indústria” e “outros serviços” houve, respectivamente, um incremento de 120,15% e 50,25%, com as taxas passando de 10,24% para 22,55% e de 11,65% para 17,51%.

A análise para “pessoas físicas” revela três períodos bem distintos: o primeiro, de junho de 1994 até março de 1996, apresenta uma elevação acelerada dos créditos inadimplentes; o segundo, de março de 1996 até outubro de 1996, caracteriza-se por uma brusca redução nos referidos créditos e, por fim, o terceiro período, a partir de outubro de 1996, com a estabilização da taxa de inadimplência em patamares próximos a 15%.

Os créditos inadimplentes na área de “habitação” mantiveram-se em níveis aceitáveis, enquanto o “setor rural”, que historicamente é lembrado como campeão da inadimplência, passa a ocupar posição secundária no ranking das atividades com maiores problemas junto ao sistema financeiro.

Figura 3 (a). Inadimplência no Setor Industrial

Figura 3 (b). Inadimplência no Setor Comercial

Figura 3 (c). Inadimplência no Setor Rural

Figura 3 (d). Inadimplência em Outros Serviços

Figura 3 (e). Inadimplência de Pessoas Físicas

Figura 3 (f). Inadimplência Habitação

Figura 3. Inadimplência nos Empréstimos do Setor Financeiro por Atividade

Fonte: Boletim do Banco Central do Brasil - Abril/1998

Obs.: Não inclui créditos em liquidação de duas instituições financeiras.

Em análise superficial sobre os fatores que levaram ao atual estágio de inadimplência nos financiamentos concedidos pelo sistema financeiro ao setor privado, pode-se destacar: (i) diminuição substancial das “*receitas inflacionárias*” nas atividades não-bancárias;⁴ (ii) elevação dos níveis de desemprego; (iii) baixa sensibilidade dos tomadores às taxas de juros;⁵ (iv) critérios inadequados de análise de crédito; e, (v) legislação leniente.

Dessa forma, com o aumento persistente dos níveis de inadimplência e a consequente redução dos valores monetários recebidos pelos bancos, a redução dos saldos de operações de crédito parece ser uma questão de tempo, pois o sistema financeiro tem interesse na concessão de crédito porque o mesmo ainda é rentável. É possível que no futuro a inadimplência barre esse crescimento.

A importância da inadimplência, segundo Araújo (1996), pode ser também analisada pelo custo de transação de um empréstimo. Considerando o custo de captação de fundos (F), o custo de administração dos créditos concedidos (A) e o prêmio de risco (PR) proveniente da inadimplência (I), todos em valores percentuais do total contratado, temos a seguinte equação para o custo de emprestar (CE):

$$CE = F + A + \frac{I(1 + F + A)}{1 - I} \quad (1)$$

onde $0 \leq I \leq 1$ e o termo $\frac{I(1 + F + A)}{1 - I}$ é o prêmio de risco.⁶

⁴ O exemplo clássico da redução dos ganhos com a inflação no setor não-bancário são os supermercados. As empresas desse segmento, no passado, trabalhavam com recebimento à vista, pagamento à prazo e giro rápido nos estoques, condições que lhes garantiram, durante um longo período de tempo, volume significativo de recursos aplicados no mercado financeiro.

⁵ Itens relacionados, principalmente, às pessoas físicas e que pelo efeito “*bola de neve*” atinge todas as atividades econômicas, ou seja, o cidadão não tem condições de quitar o seu crédito junto ao comércio, que por sua vez tem dificuldades de realizar os pagamentos devidos às instituições financeiras, indústrias, atacadistas, etc.

⁶ Esse componente, ainda segundo Araújo (1996), origina-se das perdas associadas à inadimplência. Com 0% de inadimplência o custo de emprestar é dado por F+A. Entretanto, se parcela do crédito

Segundo Baker et al. (1979), citados por Araújo (1996), é importante frisar que quando a taxa de inadimplência cresce, sua contribuição para os custos de emprestar aumenta mais que proporcionalmente, quando comparada aos efeitos dos outros componentes. Denomina-se esse fenômeno como o efeito multiplicativo da taxa de inadimplência, ou seja, a inadimplência resulta também em perdas nos custos de captação e administração da atividade de emprestar, e não somente no principal emprestado.

Ao proceder-se à derivada parcial da eq. (1) em relação aos seus componentes, tem-se:

$$\frac{\partial CE}{\partial F} = 1 + \frac{I}{1-I} = \frac{1}{(1-I)} \quad (2)$$

$$\frac{\partial CE}{\partial A} = 1 + \frac{I}{1-I} = \frac{1}{(1-I)} \quad (3)$$

$$\frac{\partial CE}{\partial I} = \frac{(1+F+A)}{(1-I)^2} \quad (4)$$

Como $1+F+A > 1-I$, tem-se:

$$\frac{\partial CE}{\partial I} > \frac{\partial CE}{\partial F} = \frac{\partial CE}{\partial A} \quad (5)$$

Sintetizando, como as variáveis estão expressas em percentagem, as derivadas percentuais representam as elasticidades dos componentes da equação 1. Desta forma, observa-se na relação exposta em (5) que o efeito na

concedido não é quitada, o tomador não está saldando também a parte dos custos administrativos e de captação atribuídos aos empréstimos não quitados.

redução de 1% na inadimplência é superior aos efeitos proporcionados pela queda de 1% em qualquer um dos outros componentes do custo de emprestar.

2.2. O "Estado da Arte" no Gerenciamento de Crédito

A preocupação quanto ao risco na concessão de empréstimos é de longa data. Já na década de 30, Keynes (1990, p. 121) advertia em relação aos riscos a que incorria o emprestador. Keynes chamava atenção para os riscos advindos de uma “contingência moral, isto é, falta voluntária ou qualquer outro meio, talvez lícito, para fugir ao cumprimento da obrigação, ou à possível insuficiência da margem de segurança, isto é, o não cumprimento involuntário causado por uma expectativa malograda”.

Keynes (1990, p. 121) acrescenta ainda um segundo motivo de risco, qual seja, “a possibilidade de uma variação desfavorável no valor do padrão monetário, tornando o empréstimo em dinheiro menos seguro, à medida da depreciação, do que um ativo real, embora todo ou maior parte deste risco já deveria ter-se refletido, e, portanto, sido incorporado, nos preços dos bens duráveis.”

Contudo, o que realmente faz-se interessante destacar, no âmbito desse trabalho, é a afirmação feita por Keynes (1990, p. 121) acerca do risco do tomador do empréstimo, que pode estender-se perfeitamente ao risco do emprestador:

“... risco é, em certo sentido, um custo real, apesar de ser suscetível de diminuição, pela média de sua distribuição, ou por efeito de uma exatidão maior nas previsões”.

No tocante ao desenvolvimento de técnicas quantitativas para modelagem dos problemas de crédito, os trabalhos pioneiros, segundo Orgler (1970) e Capon (1982), são Myers (1962) e Myers & Forgy (1963). O instrumental utilizado nesses casos foi a análise discriminante. Basicamente, procurava-se através das características individuais dos clientes, obtidas em questionários fornecidos a instituições de crédito americanas, separá-los em dois grupos distintos. Dentro do primeiro grupo ficavam aqueles cujas pontuações situavam-se acima do nível pré-estabelecido como nível de corte, recebendo a aprovação no pedido de crédito. No segundo grupo, encontravam-se os clientes com pontuação inferior ao mínimo exigido, os quais tinham suas propostas rejeitadas.⁷

Os trabalhos voltados para a previsão de inadimplência de pessoas físicas junto ao setor financeiro predominam na literatura econômica especializada desde a década de 60. Estudos relacionados ao atraso no pagamento de hipotecas e faturas de cartão de crédito são encontrados com certa frequência em periódicos da área financeira. Como exemplo, pode-se citar Koh & Tan (1994), Smith et al. (1996) e Pinder (1996).

As críticas aos modelos de previsão de inadimplência para pessoas físicas baseiam-se, principalmente, nas características individuais consideradas relevantes para a explicação dos problemas com os recebimentos dos créditos. Nos primeiros modelos desenvolvidos pela *Fair, Isaac & Company*, empresa pioneira na utilização de sistemas de crédito por pontuação, características como sexo, raça, cor dos cabelos, altura, peso, primeiro e último dígitos da

⁷ Para maiores detalhes, veja: MYERS, James H. (1962). Numerical Scoring Systems for Retail Credit Evaluation. **The Credit World**, 50 (April), 5-7, e MYERS, James H.; FORGY, Edward W (1963). The Development of Numerical Credit Evaluation Systems. **Journal of American Statistical Association**, 58 (September), 799-806.

carteira de seguro social, dentre outros, eram considerados estatisticamente relevantes para a previsão acerca da performance dos pagamentos.⁸

Segundo Lawrence e Arshadi (1995), surpreendentemente - dada a importância crescente do problema com créditos inadimplentes - existe uma carência de literatura contendo estudos empíricos sobre a resolução de problemas relacionados ao gerenciamento de créditos comerciais em instituições bancárias. Os autores acreditam que a deficiência dos dados e/ou a dificuldade de acesso aos mesmos sejam a causa principal dessa situação, dada a natureza sensível dos arquivos com informações sobre créditos e a existência de legislação federal referente ao sigilo bancário.

Orgler (1970) também indica alguns problemas limitantes aos trabalhos relacionados com os empréstimos do setor financeiro ao comércio. A primeira dificuldade reside na inexistência de uma vasta população homogênea, fato presente nos créditos pessoais, não havendo portanto, dados suficientes para um estudo estatisticamente significativo. A segunda dificuldade diz respeito a substanciais variações existentes entre os créditos comerciais em relação ao tamanho das empresas, a garantias oferecidas e formas de pagamento, sendo que no caso dos empréstimos pessoais há uma relativa uniformidade nesses termos. E, finalmente, o autor discorre sobre as dificuldades em obter dados financeiros atualizados das pequenas empresas comerciais que recorrem a linhas de crédito.

A existência de trabalhos analisando o desempenho das empresas comerciais começa com modelos de previsão de insolvência e datam do início da década de 30. Esses trabalhos, utilizando dados financeiros publicados pelas empresas, objetivavam, em sua maioria, detectar os índices mais significativos na explicação do processo de falência de algumas empresas

⁸ Capon (1982), p. 85.

americanas. Contudo, apenas no final da década de 60, mais precisamente com Altman (1968), surgiu um modelo baseado em instrumentos estatísticos mais evoluídos, qual seja, a análise discriminante múltipla.

Nos anos seguintes vários trabalhos foram publicados nesta mesma linha de pesquisa, ou seja, previsão de insolvência utilizando análise discriminante. O diferencial residia apenas no período de tempo e na região escolhida para o estudo. Para o Brasil se destacam os estudos de Kanitz (1978) e de Elizabetsky (1976), que apresentam um modelo para decisão de crédito em bancos comerciais e utilizam dados de empresas da indústria de confecções⁹; o de Matias (1978)¹⁰ e do próprio Altman (1979).¹¹

Recentemente, os estudos referentes aos problemas com gerenciamento de crédito voltaram a se destacar na literatura econômica. Os artigos, em sua maioria, utilizam modelos de resposta qualitativa, com destaque maior para os modelos logísticos e modelos baseados em redes neurais.

Comparando as técnicas de análise discriminante e próbite, Grabloesky e Talley (1981), chegam à conclusão de que para a classificação de propostas de obtenção de empréstimos com risco reduzido, as duas técnicas são semelhantes, enquanto para propostas com riscos mais elevados o próbite apresenta uma taxa de classificação correta maior.

⁹ Elizabetsky, Roberto (1976). **Um modelo matemático para decisão de crédito no banco comercial.** (trabalho apresentado ao Depto. de Engenharia de Produção da Escola Politécnica da USP).

¹⁰ Matias, Alberto B (1978). **Contribuição às Técnicas de Análise Financeira: Um Modelo de Concessão de Crédito.** (Trabalho apresentado ao Depto. de Administração da Faculdade de Economia e Administração da USP).

¹¹ Altman, E.; Tara, K. N.; Dias, L. M. R. Previsão de Problemas Financeiros em Empresas. **Revista de Administração de Empresas**, v. 19, n. 1, 1979.

Unger (1988) analisa as variáveis que afetam a capacidade de crédito dos solicitantes de recursos. As variáveis financeiras escolhidas são: liquidez, solvência, lucratividade, garantias e capacidade de pagamento. O autor também faz considerações a respeito de características pessoais e gerenciais dos clientes.

Leonard & Banks (1994) apresentam um modelo baseado em redes neurais para identificar as variáveis significativas no processo de previsão de inadimplência em créditos pessoais e estimam a redução das perdas a que incorrem os emprestadores quando da adoção de modelos de pontuação do crédito.

Lawrence e Arshadi (1995) apresentam um estudo acerca das medidas adotadas por 52 bancos¹² em 25 estados americanos para solucionar os problemas com os créditos inadimplentes. As quatro alternativas passíveis de serem seguidas pelos emprestadores são: (i) executar judicialmente o crédito; (ii) requerer sua falência, numa ação conjunta com o tomador do empréstimo; (iii) renegociação, com mudanças nas taxas de juros e nos prazos; e (iv) fornecer um crédito adicional para o tomador sair do “status” de inadimplente.¹³

Em um outro artigo, Angbazo (1997) mostra-nos a relação entre a inadimplência, cuja importância é ressaltada pelo autor nos recentes problemas ocorridos com o gerenciamento de crédito, e a margem de lucro líquida nos bancos comerciais americanos. O resultado do modelo adotado por Angbazo indica que a inadimplência afeta sobremaneira as margens líquidas de lucro nos bancos caracterizados como “*money-center banks*” e nos bancos considerados pequenos, enquanto para os bancos regionais e supra regionais

¹² Os bancos foram escolhidos aleatoriamente de uma população composta por bancos de pequeno e médio portes (ativos menores que US\$ 1 bilhão).

¹³ Os resultados obtidos com a utilização de um modelo Lógite multinomial confirmam que a alternativa escolhida dependerá da combinação de variáveis referentes aos emprestadores e aos tomadores.

as evidências não apontam a inadimplência como fator significativo nas margens de lucro.¹⁴

Os modelos para classificação do grau de risco do crédito encontraram um campo extremamente fértil no setor agrícola. Historicamente, esse setor caracteriza-se como problemático nos pagamentos de seus empréstimos, apresentando elevadas taxas de inadimplência.

Segundo Araújo (1996), destacam-se os estudos de Bauer & Jordan (1971), Dunn & Frey (1976), Hardy e Weed (1980) - todos utilizando análise discriminante - e ainda Fischer & Moore (1986), Miller & LaDue (1989), Turvey & Brown (1990) e Knopf & Schoney (1992) que utilizam modelos logísticos.¹⁵

O próprio Araújo (1996), utilizando dados do Banco do Brasil, desenvolveu um modelo logístico para avaliação e classificação do grau de risco das propostas apresentadas por cooperativas agropecuárias. O modelo estimado pelo autor classificou corretamente 78,5% das observações, sendo especificado com as seguintes variáveis: liquidez corrente, endividamento interno, passivo oneroso e saldo de tesouraria.

¹⁴ No caso dos pequenos bancos e dos bancos regionais e supra regionais as evidências também apontam para a taxa de juros de risco como fator explicativo para as margens de lucro.

¹⁵ Araújo (1996), op. cit., p. 15-20.

3. REFERENCIAL TEÓRICO

Um modelo de previsão de inadimplência (também conhecido, na prática, como modelo de escoragem de crédito) deve, basicamente, caracterizar-se por: (i) combinar os fatores mais importantes associados à possibilidade de ocorrência de tal fenômeno, (ii) determinar o inter-relacionamento entre esses fatores e (iii), por último, quantificá-los. Esse modelo deve ser capaz de produzir um “escore final” diretamente relacionado à probabilidade de inadimplência, ou seja, quanto maior for o “escore final” maior será o risco de perda (inadimplência) nas operações de crédito analisadas.

Neste trabalho utilizar-se-ão quatro métodos distintos na construção dos modelos de previsão de inadimplência para os financiamentos concedidos pelo BANDES ao segmento de hotelaria (financiamentos de longo prazo). Os três primeiros métodos estão baseados no conceito de “variáveis qualitativas” e o último, no conceito de “entropia”.

Os modelos que utilizam o conceito de resposta qualitativa, são capazes de quantificar (ponderar) a participação de cada uma das variáveis pré-determinadas no fenômeno da inadimplência, enquanto o modelo baseado no conceito de entropia tem o objetivo de estabelecer uma ordenação entre variáveis pré-determinadas. As características básicas da carteira de operações do BANDES e dos modelos propostos estão descritas a seguir.

3.1. O BANCO DE DESENVOLVIMENTO DO ESPÍRITO SANTO S/A - BANDES

Vinculado à Secretaria da Fazenda do Estado do Espírito Santo, o Banco de Desenvolvimento do Espírito Santo S/A (BANDES) é a instituição financeira responsável - em conjunto com o Grupo Executivo de Recuperação Econômica do Espírito Santo (GERES) - pela execução da política de incentivos fiscais e creditícios do Estado do Espírito Santo.

O total de investimentos apoiados pelo BANDES nos últimos anos, assim como o número de operações de crédito e de empregos diretos gerados, podem ser visualizados na tabela 2. A elevação no número de operações de crédito a partir de 1995 explica-se pelas medidas adotadas pelo banco para viabilizar empréstimos a segmentos sociais que sempre tiveram grande dificuldade em obtê-lo junto a instituições bancárias, sejam públicas ou privadas. Os “segmentos” beneficiados pelas medidas foram: micro e pequenas empresas, produção familiar rural, jovens empresários (recém formados), estudantes de pós-graduação, pesca artesanal e setor informal.

Tabela 2. BANDES: Investimentos Aprovados: 1991/1996

Ano	Número de Operações	Investimentos (US\$)	Empregos Gerados
1991	467	43.335.198	2.553
1992	517	167.775.764	4.993
1993	521	133.946.136	3.779
1994	342	138.301.591	2.792
1995	428	70.227.685	1.603
1996	1.212	105.632.746	3.919

Fonte: BANDES

O apoio a tais segmentos, de grande importância social, só foi possível por se tratar de instituição pública, na qual, muitas vezes, o retorno privado é preterido em função do retorno social. Ademais, esses empréstimos atuaram

fortemente na diversificação da carteira de empréstimos do banco, diluindo os riscos inerentes à mesma.

As principais fontes de recursos utilizadas pelo BANDES para o financiamento dessas operações, além dos próprios recursos do banco, são os fundos dos quais é o agente técnico e financeiro, quais sejam, o Fundo de Desenvolvimento das Atividades Portuárias (FUNDAP)¹⁶ e o Fundo de Recuperação Econômica do Espírito Santo (FUNRES)¹⁷. Em nível federal, o BANDES mobiliza e repassa recursos de fontes como: Banco Nacional de Desenvolvimento Econômico e Social (BNDES); Financiadora do Estudo e Projetos (FINEP), Agência Especial de Financiamento (FINAME), Empresa Brasileira de Turismo (EMBRATUR) e Caixa Econômica Federal (CEF).

A participação dos fundos estaduais no total de investimentos aprovados pelo BANDES é significativa, com exceção de 1994, onde predominaram as fontes federais, conforme demonstrado na figura 4. Dentre as fontes de recursos utilizadas para investimentos, deve-se ressaltar que os “Recursos Próprios” correspondem às inversões feitas pelos grupos empresariais responsáveis pelos empreendimentos.

¹⁶ Criado pela Lei 2508, de 22 de Maio de 1970.

¹⁷ Criado pelo decreto-lei N° 880, de 18 de Setembro de 1969, o mesmo que instituiu os fundos de desenvolvimento das regiões Norte e Nordeste do país.

Figura 4. BANDES: Investimentos Aprovados por Fonte de Financiamento

Fonte: BANDES

Na figura 5 mostra-se o montante de investimentos aprovados por objetivo nos últimos anos, de onde pode-se observar que a maioria absoluta dos investimentos está voltada para a implantação de empresas.

Quanto aos investimentos por atividade, figura 6, os destaques ficam por conta dos setores secundário e terciário, com o último sendo responsável pelo maior aporte de recursos em projetos aprovados pelo BANDES em 1996. Dadas as características atuais da economia do Espírito Santo, acredita-se que os investimentos no setor terciário deverão permanecer em patamares elevados nos próximos anos. Dentro desse setor, subdividido em comércio e serviços, destacam-se os segmentos de armazenagem, serviços portuários e hotelaria.

Figura 5. BANDES: Investimentos Aprovados por Objetivo

Fonte: BANDES

Figura 6. BANDES: Investimentos Aprovados por Atividade

Fonte: BANDES

3.1.1. A INADIMPLÊNCIA NO BANDES

Atualmente o BANDES possui um saldo de operações de crédito no valor aproximado de R\$ 120 milhões, com os créditos em atraso e em liquidação, ou seja, a inadimplência, representando algo em torno de 18% desse valor¹⁸. A tabela 3 mostra a inadimplência média nas instituições financeiras nacionais e a média de inadimplência nos empréstimos do BANDES, tais médias podem ser comparadas com os níveis recentemente observados no BANDES.

Tabela 3. Níveis de Inadimplência (%): Média Nacional e BANDES

Inadimplência	Dez/95		Jun/96		Dez/96
	Média	Bandes	Média	Bandes	Bandes
Total	9,00	15,80	10,00	22,00	18,40
BANDES	2,00	14,50	2,00	20,00	13,80

Fonte: BANDES

Como se pode observar, o BANDES apresenta níveis de inadimplência em valores superiores à média nacional, tanto em relação ao total de saldo de operações de crédito, quanto ao total repassado pelo BANDES.¹⁹ Além dos problemas citados anteriormente por consequência da inadimplência, no caso específico do BANDES há alguns agravantes a serem comentados.

¹⁸ Para efeito de comparação, observe os dados relativos a Superintendência do Banco do Brasil no Espírito Santo, cujo saldo de operações de crédito alcança valores próximos a R\$ 1,3 bilhão. A inadimplência chega a 18%, algo próximo a R\$ 240 milhões, segmentados da seguinte forma: indústria 62%, comércio 13%, agricultura 9%, pecuária 2%, câmbio e cheque-ouro 14%.

¹⁹ A comparação dos níveis de inadimplência entre diferentes tipos de instituições financeiras apresenta um viés relacionado às características da cada instituição. No caso dos bancos comerciais, que exigem "reciprocidade" para a liberação dos recursos, trabalha-se evidentemente com perdas menores, seja pela própria existência de recursos do tomador aplicados na instituição, seja pela comercialização de produtos (seguros, por exemplo). No caso dos bancos de desenvolvimento, as exigências de reciprocidade (quando existem) são de menor intensidade.

O primeiro problema diz respeito à redução e/ou indisponibilidade da aplicação dos recursos provenientes do BNDES e das outras fontes federais²⁰. O segundo, atrelado ao problema anterior, é que o sistema de crédito estadual assumiu importância significativa no momento de intensificação das disputas regionais por investimentos, não podendo, portanto, o mesmo apresentar-se fragilizado e incapacitado na sua função de indução do desenvolvimento econômico e social do Estado.

Outro problema relacionado com a inadimplência diz respeito ao próprio desempenho econômico-financeiro do BANDES. Pelo terceiro ano consecutivo o banco apresenta prejuízo em seu balanço. Os prejuízos em 1994, 1995 e 1996 foram, respectivamente, R\$ 1,3 milhão, R\$ 4,3 milhões e R\$ 13 milhões. No ano de 1996 a recuperação de créditos baixados em prejuízo foi da ordem de R\$ 5 milhões, enquanto a provisão de créditos duvidosos foi algo em torno de R\$ 1,5 milhão. Esses prejuízos reduzem sobremaneira a disponibilidade de recursos próprios, além de prejudicar a captação de outras fontes de recursos, inviabilizando o apoio, por parte do BANDES, a novos investimentos no Espírito Santo.

Em análise mais detalhada sobre a inadimplência do BANDES, procurou-se identificar as atividades com maiores problemas junto ao banco, conforme mostra a tabela 4. Algumas atividades, apesar de apresentarem elevados índices de inadimplência, não foram citadas, seja pela inexpressiva participação na carteira de crédito do banco, seja por questões relacionadas ao sigilo bancário.²¹

²⁰ O nível de inadimplência das instituições financeiras é um dos índices utilizados pelo BNDES para a definição dos recursos a serem repassados pelas mesmas. No caso do BANDES, o limite de repasse de recursos do BNDES caiu de R\$ 29 milhões em Junho de 1996 para R\$ 12 milhões em Dezembro de 1996. Somente nos últimos meses, após grande pressão política, houve a elevação do limite para o mesmo patamar de Junho/96.

²¹ Nesse caso incluem-se atividades em que o número de firmas é extremamente reduzido, possibilitando o reconhecimento da(s) empresa(s) em questão.

Tabela 4. BANDES: Taxa de Inadimplência por Atividade (Base: Março/1997)

Atividade	Número de Operações	Saldo Devedor Total (R\$ Mil)	Saldo Operação de Crédito (R\$ Mil)	Taxa de Inadimplência (%)
Mármore e Granito	85	29.666	27.167	12,15
Hotelaria	97	31.888	22.378	15,24
Produtos Alimentares ¹	23	30.597	20.650	4,30
Confeção	118	11.145	9.062	41,18
Armazéns Gerais	12	9.187	8.570	5,20
Serviços de Saúde	64	8.076	5.160	3,15

(1) Massas Alimentícias, Frigoríficos e Refinação e Moagem de Açúcar.
 Fonte: BANDES

Uma consideração deve ser feita ao item “Saldo Devedor Total” na tabela 4. Observe, por exemplo, a atividade “produtos alimentares”. Apesar de apresentar um pequeno índice de inadimplência, cerca de 4,3%, essa atividade já teve créditos baixados em prejuízo no valor aproximado de R\$ 9,9 milhões. Portanto, se considerarmos o total de créditos não recebidos nessa atividade, o “*índice de inadimplência*” nos financiamentos chega a 35,41%. Fazendo o mesmo para “hotelaria” e “serviços de saúde” obtemos, respectivamente, os índices de 40,52% e 38,12%.

Para amenizar essa situação, e diminuir os riscos inerentes às operações de crédito, o BANDES vem tomando algumas iniciativas na avaliação e na concessão de financiamentos. A primeira é a maior diversificação de suas operações em termos de atividade, tamanho, objetivo e endividamento dos projetos. A segunda trata da confecção de uma matriz de riscos, cujo objetivo é estimar o risco associado a cada novo projeto, levando-se em conta variáveis micro e macroeconômicas.

Tais medidas, de caráter preventivo, devem apresentar resultados positivos num horizonte de médio e longo prazos. Para o curto prazo, o Banco

adotou duas estratégias básicas: (i) a terceirização da cobrança, visando, principalmente, o recebimento dos créditos em atraso e em liquidação (para os créditos em prejuízo, o banco vem estudando, caso a caso, a possibilidade de execução judicial); e, (ii) a securitização dos créditos inadimplentes e em prejuízo.²²

Portanto, nesta seção procurou-se demonstrar a elevada inadimplência verificada nos financiamentos do BANDES, assim como os problemas advindos dessa situação. Tais problemas justificam a utilização de novos métodos para análise e concessão de crédito que, em conjunto com as medidas já adotadas, objetivem a redução dos níveis de inadimplência.

3.2. MODELOS DE RESPOSTA QUALITATIVA

Os modelos de resposta qualitativa são modelos de regressão cujas variáveis dependentes assumem valores discretos. Em nosso caso específico, trabalha-se com modelos em que tais variáveis podem assumir dois valores, quais sejam, 0 ou 1. O valor “Zero” será atribuído para a não ocorrência da inadimplência, enquanto o valor “Um” será atribuído para a ocorrência da inadimplência.

Segundo Amemiya (1985), os modelos de resposta qualitativa binária univariada,²³ os quais utilizaremos em nosso trabalho, são definidos da seguinte forma:

$$P(y_i = 1) = F(x_i' \beta_0) \quad i = 1, 2, \dots, n \quad (6)$$

²² Nesse caso, entende-se por securitização o processo de repasse a terceiros de créditos considerados “problemáticos”. O deságio em relação ao valor total do crédito será diretamente proporcional ao risco de não recebimento do mesmo.

²³ Os modelos multivariados e multinomiais podem ser vistos com detalhes em Amemiya (1985), Greene (1993), Judge et al. (1988).

onde y_i são as variáveis aleatórias binárias independentes com valor “Zero” ou “Um”, x_i é um vetor de tamanho k , com constantes conhecidas, β_0 é um vetor de parâmetros desconhecidos, também de tamanho k , e F é uma certa função conhecida.

Os métodos de estimação e inferência a serem utilizados em nosso trabalho compreendem os seguintes modelos, com as respectivas formas funcionais conhecidas da função F :

- Modelo de Probabilidade Linear

$$F(x) = x \quad (7)$$

- Modelo Probabilístico (*Próbite*)

$$F(x) = \Phi(x) \equiv \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{t^2}{2}\right)\right] dt \quad (8)$$

- Modelo Logístico (*Lógite*)

$$F(x) = \Lambda(x) \equiv \frac{e^x}{1 + e^x} \quad (9)$$

3.2.1. MODELO UTILIZANDO PROBABILIDADE LINEAR²⁴

Considere o seguinte modelo simplificado:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \quad (10)$$

onde Y_i pode assumir somente dois valores, ou valor 0, ou valor 1, e u_i representa o termo estocástico.

Modelos como o apresentado na eq. (10), em que a variável dependente Y_i é função linear da variável explicativa X_i , são conhecidos como modelos de probabilidade linear desde que $E(Y_i | X_i)$, esperança condicional de Y_i dado X_i , possa ser interpretada como a *probabilidade condicional* de que o evento ocorrerá dado X_i , isto é, $\Pr(Y_i = 1 | X_i)$.

Assumindo que $E(u_i) = 0$, de forma a obter-se estimadores não tendenciosos, obtêm-se a eq. (11):

$$E(Y_i | X_i) = \beta_1 + \beta_2 X_i \quad (11)$$

Considerando que o evento $(Y_i=1|X_i)$ ocorre com probabilidade P_i e, conseqüentemente, $(Y_i=0|X_i)$ com probabilidade $1-P_i$, a variável Y_i tem a seguinte distribuição:

$E(Y_i X_i)$	Probabilidade
0	$1 - P_i$
1	P_i
Total	1

²⁴ Para maiores detalhes sobre modelos de probabilidade linear consultar Gujarati (1995) e Greene (1993).

Desse modo, usando a definição de esperança matemática, tem-se:

$$E(Y_i | X_i) = 0(1 - P_i) + 1(P_i)$$

$$E(Y_i | X_i) = P_i \quad (12)$$

Comparando a eq. (11) com a eq. (12), pode-se rescrevê-las como:

$$E(Y_i | X_i) = \beta_1 + \beta_2 X_i = P_i \quad (13)$$

onde a esperança condicional, $E(Y_i | X_i)$, do modelo mostrado na eq. (10) pode, efetivamente, ser interpretada como a probabilidade condicional de Y_i .

3.2.1.1. PROBLEMAS COM MODELOS DE PROBABILIDADE LINEAR.

A estimação desses modelos pelo método de mínimos quadrados ordinários (MQO) apresenta alguns problemas, a saber: (i) a não normalidade do termo estocástico u_i ; (ii) a variância do termo estocástico u_i não ser constante (heterocedasticia); (iii) o valor de R^2 ser questionável como medida de qualidade do ajuste; (iv) o não cumprimento da condição $0 \leq E(Y_i | X_i) \leq 1$; e, (v) assumir-se que $P_i (Y_i = 1 | X)$ cresce linearmente com X , isto é, o efeito marginal de X permanece constante.

O problema da não normalidade de u_i não é crítico, uma vez que o objetivo é estimar um ponto, sendo desnecessário assumir a hipótese de normalidade. Ademais, com grandes amostras, pode-se mostrar que os estimadores obtidos via MQO tendem a ser normalmente distribuídos.²⁵ Assim, as inferências estatísticas sobre os modelos de probabilidade linear seguirão

²⁵ A prova desse argumento baseia-se no Teorema do Limite Central e, segundo Gujarati (1995), pode ser encontrada em E. Malinvaud, *Statistical Methods of Econometrics*, Rand McNally & Company, Chicago, 1966, p. 195-197.

os procedimentos usuais do método de MQO, assumindo a hipótese de normalidade.

Quanto a heterocedasticia, a mesma pode ser corrigida ponderando as observações pelas variâncias, ou seja, utilizando-se o método de mínimos quadrados generalizados. Enquanto o uso do coeficiente de determinação (R^2) como um resumo estatístico pode ser descartado em modelos com variáveis dependentes qualitativas.²⁶

Visto que $E(Y_i | X)$ mede a probabilidade condicional do evento Y ocorrer dado X , a mesma necessariamente deve estar entre 0 e 1. Embora seja verdade a priori, não existem garantias de que o estimador de $E(Y_i | X)$ respeitará essa condição, sendo esse um problema crítico com a estimação via MQO para modelos de probabilidade linear.

E, finalmente, a hipótese de crescimento linear de P_i ($Y_i = 1 | X$) em relação a X mostra-se fortemente equivocada. Na realidade, conforme afirma Gujarati (1995), espera-se que a relação entre P_i e X seja não linear.²⁷

²⁶ Geralmente trabalha-se com modelos de probabilidade linear restritos (ou truncados). Esse procedimento visa garantir que o modelo estimado fique dentro da banda 0 - 1. Como resultado, o convencional R^2 fica muito abaixo de 1 nesses modelos.

²⁷ Para maiores detalhes veja Gujarati (1995), p. 552-4.

Assim, para satisfazer a condição de P_i entre 0 e 1 e garantir sua não-linearidade com X pode-se utilizar uma função de distribuição cumulativa, conforme apresentada na figura 7.

Figura 7. Função de Distribuição Cumulativa

Por razões históricas, e também por razões práticas, as funções de distribuição cumulativa usualmente escolhidas para representar os modelos de resposta 0 e 1 são as funções logística e normal, a primeira originando o modelo *Lógite* e a segunda o modelo *Próbite*.

3.2.2. MODELOS LÓGITE E PRÓBITE²⁸

Considere a variável Y_i^* definida pela seguinte regressão:

$$Y_i^* = \beta' X_i + u_i \quad (14)$$

Na prática, Y_i^* não é observável. O que se observa é a variável *dummy* Y definida por:

$$\begin{aligned} Y &= 1 && \text{se } Y_i^* > 0 \\ Y &= 0 && \text{nos demais casos} \end{aligned} \quad (15)$$

²⁸ Para maiores detalhes sobre o assunto consultar Gujarati (1995) e em Maddala (1983).

Nessa formulação $\beta'X_i$ não é mais $E(Y_i|X_i)$, como no modelo de probabilidade linear. Agora, $\beta'X_i$ é definido como $E(Y_i^*|X_i)$. Das eq. (14) e (15), obtêm-se

$$P_i = \Pr(Y_i = 1) = \Pr(Y_i^* > 0) = \Pr(\beta'X_i + u_i > 0) = \Pr(u_i > -\beta'X_i)$$

$$P_i = \Pr(Y_i = 1) = 1 - F(-\beta'X_i) \quad (16)$$

onde F é a função de distribuição cumulativa para u .

No caso, os valores observados de Y são sempre realizações de um processo binomial com probabilidades dadas pela eq. (16) e variando de ensaio para ensaio (dependendo de X_i). Portanto, a função de verossimilhança é:

$$L = \prod_{Y_i=0} F(-\beta'X_i) \prod_{Y_i=1} [1 - F(-\beta'X_i)] \quad (17)$$

A forma funcional para F na eq. (17) dependerá das hipóteses acerca de u_i na eq. (5). Se a distribuição de u_i for a logística, tem-se o modelo Lógite. Nesse caso:

$$F(-\beta'X_i) = \frac{e^{-\beta'X_i}}{1 + e^{-\beta'X_i}} = \frac{1}{1 + e^{\beta'X_i}}$$

Portanto,

$$P_i = \Pr(Y_i = 1) = 1 - F(-\beta'X_i) = \frac{e^{\beta'X_i}}{1 + e^{\beta'X_i}} \quad (18)$$

Em (18) observa-se facilmente que P_i relaciona-se de forma não linear com X_i e, quando $\beta'X_i$ varia de $-\infty$ a $+\infty$ P_i , limita-se entre 0 e 1. Assim, com a adoção da função de distribuição logística, satisfaz-se as duas condições requeridas anteriormente.

No caso acima, diz-se existir uma expressão na forma fechada para F , pois não envolve-se explicitamente integrais. Contudo, nem todas as

distribuições permitem tal expressão na forma fechada. Por exemplo, no modelo Próbite assume-se que u_i tem distribuição normal com média igual a zero e variância igual a σ^2 . Nesse caso:

$$F(-\beta'X) = \int_{-\infty}^{\frac{-\beta'X}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left[-\left(\frac{t^2}{2}\right)\right] dt \quad (19)$$

onde inicialmente pode-se considerar σ igual a 1, uma vez que podemos estimar apenas β/σ , e não β e σ separadamente.

Novamente, agora na eq. (19), observa-se que P_i relaciona-se de forma não linear com X_i e, quando $\beta'X_i$ varia de $-\infty$ a $+\infty$ P_i , limita-se entre 0 e 1. Assim, com a adoção da função de distribuição normal, satisfaz-se também as duas condições requeridas anteriormente.

3.2.2.1. ESTIMAÇÃO DOS PARÂMETROS

Ao contrário do modelo de probabilidade linear, no qual as estimativas dos parâmetros são relativamente fáceis de serem realizadas (através de MQO), os modelos Lógite e Próbite exigem métodos mais complexos de estimação.

Nos casos dos modelos Lógite e Próbite as estimações dos parâmetros são realizadas pelo método da máxima verossimilhança. Especificamente para o Lógite, a eq. (17) pode ser reescrita como:

$$L = \prod_{i=1}^n \left(\frac{1}{1 + e^{\beta'X_i}} \right)^{1-Y_i} \left(\frac{e^{\beta'X_i}}{1 + e^{\beta'X_i}} \right)^{Y_i} \quad (20)$$

ou

$$L = \frac{e^{\beta' \sum_{i=1}^n X_i Y_i}}{\prod_{i=1}^n [1 + e^{\beta' X_i}]} \quad (21)$$

Colocando na forma logarítmica e definindo $t^* = \sum_{i=1}^n X_i Y_i$, têm-se:

$$\log L = \beta' t^* - \sum_{i=1}^n \log[1 + e^{\beta' X_i}] \quad (22)$$

A estimativa dos β 's pelo método da máxima verossimilhança, no qual maximiza-se $\log L$ em relação a β e iguala-se a zero, ou seja, $\partial \log L / \partial \beta = 0$, é dada por:

$$S(\beta) = - \sum_{i=1}^n \frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}} X_i + t^* = 0 \quad (23)$$

Para o modelo Próbite, a função de verossimilhança correspondente à eq. (17) é dada por:

$$L = \prod_{i=1}^n [\Phi(\beta' X_i)]^{Y_i} [1 - \Phi(\beta' X_i)]^{1-Y_i} \quad (24)$$

que na forma logarítmica fica:

$$\log L = \sum_{i=1}^n Y_i \log \Phi(\beta' X_i) + \sum_{i=1}^n (1 - Y_i) \log [1 - \Phi(\beta' X_i)] \quad (25)$$

A estimativa dos β 's pelo método da máxima verossimilhança, no qual se maximiza $\log L$ em relação a β e iguala-se a zero, ou seja, $\partial \log L / \partial \beta = 0$, é dada por:

$$S(\beta) = \sum_{i=1}^n \frac{[Y_i - \Phi(\beta' X_i)]}{\Phi(\beta' X_i) [1 - \Phi(\beta' X_i)]} \Phi(\beta' X_i) X_i = 0 \quad (26)$$

onde $\Phi(\cdot)$ e $\phi(\cdot)$ são, respectivamente, as funções de distribuição e de densidade normal.

Observa-se que as equações para a estimação dos parâmetros nos modelos Lógite e Próbite são não-lineares nos parâmetros β 's, requerendo, portanto, soluções pelo uso de processos iterativos. Os algoritmos utilizados com maior frequência para a estimação desses parâmetros são o método de Newton-Raphson, o método de Score e o método de Berndt, Hall, Hall e Hausman (BHHH).²⁹

3.2.3. CONSIDERAÇÕES SOBRE MODELOS DE RESPOSTA QUALITATIVA

As considerações finais a serem feitas sobre os modelos expostos anteriormente baseiam-se em comparações quanto aos motivos de utilizar-se, ou não, determinado modelo. No caso do MPL, conforme destaca Amemiya (1985), a principal falha está no fato de a forma funcional F desse modelo não ser uma função de distribuição apropriada, uma vez que não está garantida a restrição do condicionamento da probabilidade entre 0 e 1. Esse problema pode ser corrigido ao definir-se $F = 1$ se $F(\beta'X_i) > 1$ e $F = 0$ se $F(\beta'X_i) < 0$, mas esse procedimento produz embaraços irreais nos pontos de truncagem.³⁰

Segundo Amemiya (1985), no caso do modelo Próbite, assim como outros modelos estatísticos que utilizam a distribuição normal, seu uso justifica-se pelo seu apego ao Teorema do Limite Central. Enquanto a principal justificativa para o modelo Lógite é que a função de distribuição logística é similar à função de distribuição normal, apresentando-se, contudo, numa forma bem mais simples.

²⁹ Para detalhes sobre os métodos de estimação em modelos não-lineares consulte Judge et al. (1988).

³⁰ Os modelos de probabilidade linear eram usados com frequência em aplicações econométricas, especialmente até a década de 70, pela sua simplicidade computacional.

Ainda segundo Amemiya (1981), em estudo sobre os modelos de resposta qualitativa, os modelos lógite e próbite normalmente apresentam resultados similares, embora os coeficientes estimados não sejam diretamente comparáveis. Segundo o autor, pode-se obter valores aproximados dos coeficientes da distribuição logística multiplicando os coeficientes do modelo próbite pela constante 1,6. Amemyia (1981) também sugere que os coeficientes do modelo de probabilidade linear podem ser obtidos multiplicando-se os coeficientes do modelo lógite por 0,25 (exceto para a constante, que seria obtida pela multiplicação do respectivo coeficiente por 0,25, devendo-se ainda ao valor final somar 0,5).

3.3. MODELO BASEADO NO CONCEITO DE ENTROPIA

O conceito de entropia relaciona-se diretamente ao grau de incerteza sobre determinado fenômeno. Diz-se que “a entropia da distribuição é máxima, ou seja, há um máximo de incerteza a respeito do que pode ocorrer, quando todos os possíveis eventos são igualmente prováveis, isto é, há um máximo de *desordem* no sistema”.³¹

O método a ser utilizado, nesse estudo, para a ordenação das variáveis e, conseqüentemente, o cálculo da entropia será o ID3, desenvolvido por Quinlan (1983), no qual o reconhecimento e a classificação dos padrões consistem na construção de uma “árvore discriminatória” com atributos numéricos e/ou não-numéricos (não-mensuráveis).

O procedimento padrão no método ID3 é a análise do ganho de informação ou, equivalentemente, o decréscimo da entropia, tomando-se como

base determinadas características da amostra estudada. Define-se, quantitativamente, entropia como:

$$Entropia = -p_1 \log_2 p_1 - p_0 \log_2 p_0 \quad (27)$$

onde a probabilidade p é determinada com base na frequência de ocorrência do fenômeno, com p_1 correspondendo a “ocorreu” e p_0 a “não ocorreu”.

A entropia será máxima quando p_1 for igual a p_0 , ou seja, a probabilidade de ocorrência do fenômeno 1 é 50% e, conseqüentemente, a probabilidade de ocorrência do fenômeno 2 também é 50%. Nesse caso, ter-se-ia uma entropia de 1 *bit*.³²

3.4. CONSIDERAÇÕES SOBRE O USO DE MODELOS NA PREVISÃO DA INADIMPLÊNCIA

A primeira consideração a ser feita na utilização do instrumental proposto reside na possibilidade de se conhecer as variáveis importantes para caracterizar e/ou prever uma situação de inadimplência. Conseqüentemente, configurando-se como um aspecto positivo, pode-se eliminar o fator “subjetividade” na análise dos projetos, pois sabe-se que é perfeitamente possível, dependendo do “*feeling*” do técnico e/ou equipe responsável pela análise, que o banco tome decisões diferentes em relação à concessão de crédito para uma mesma operação.

Outra consideração a ser feita refere-se à maior agilidade no processo de tomada de decisão. Além de se configurar como importante estratégia

³¹ Hoffmann, Rodolfo (1991), p. 289.

³² Quando se utilizam logaritmos de base 2 diz-se que o conjunto informativo é medido em “*bits*” e quando se utilizam logaritmos naturais diz-se que o conteúdo informativo é medido em “*nits*”. Desta forma, conclui-se que: 1 “*bit*” = 0.693 “*nit*”.

competitiva frente aos demais bancos,³³ uma vez que para o tomador do empréstimo a urgência na liberação dos recursos pode ser fundamental, a maior agilidade permitirá aos técnicos do BANDES dedicar parcela maior de seu tempo a outros assuntos relevantes que não podem ser sistematizados e/ou à análise mais detalhada dos projetos que estejam em situação indefinida quanto a probabilidade de inadimplência.

No caso das desvantagens (ou limitações) está o entendimento de que o modelo proposto não deve ser encarado como verdade única, uma vez que dificilmente seu nível de acerto será de 100%. Evidentemente que o mesmo pode apresentar erros do tipo: classificar um projeto como sendo de reduzido risco, quando na realidade não o é (erro do tipo 1), ou classificar um projeto como sendo de elevado risco, quando na realidade o risco é baixo (erro do tipo 2).

A questão do tempo também é importante. A alternativa de se considerar dados históricos para predizer comportamentos futuros pode levar a conclusões equivocadas, uma vez que tais dados estão condicionados a peculiaridades de seu tempo, os quais não necessariamente poderão ocorrer no futuro. Uma característica pode ser importante em determinada época e deixar de sê-la em períodos subsequentes.

Outros itens limitadores do uso do modelo para previsão de inadimplência são: (i) a falta de padronização dos projetos e das análises dos mesmos, o que exige atenção às indicações das variáveis especificadas no modelo; (ii) aspectos geográficos e setoriais podem ser importantes na limitação do uso de um único modelo, podendo gerar demandas por modelos mais específicos, de acordo com a necessidade do BANDES.

³³ Válido para o caso de financiamentos realizados com recursos que não sejam operacionalizados exclusivamente pelo BANDES, como por exemplo os recursos provenientes de fontes federais.

Ainda no caso específico do BANDES, pelo fato de tratar-se de um banco público com responsabilidades sobre a indução do desenvolvimento econômico no Estado do Espírito Santo, poderão existir algumas decisões de financiamento que contrariem as indicações feitas pelo modelo, motivadas por questões políticas ou por questões relacionadas ao desenvolvimento regional.

4. ANÁLISE DOS RESULTADOS

As metodologias propostas para os modelos de previsão de créditos inadimplentes podem contemplar todas as operações realizadas pelo BANDES, contudo, dada a especificidade de cada segmento econômico apoiado pelo banco, decidiu-se por realizar uma análise setorial, objetivando resultados mais precisos. O segmento escolhido, seguindo a classificação utilizada pelo BANDES, foi o de “hotéis, motéis e pousadas”, doravante denominado como segmento “hoteleiro”.

A escolha do segmento hoteleiro justifica-se, basicamente, por quatro motivos: (i) importância dos créditos inadimplentes nesse setor em relação ao total de crédito inadimplente (estima-se que o segmento de hotelaria responde por cerca de 30% da inadimplência, em valores monetários, junto ao BANDES); (ii) elevado número de financiamentos realizados; (iii) desconcentração espacial dos projetos; e, (iv) importância cada vez maior da “indústria do turismo” enquanto atividade geradora de impostos, renda e emprego.

Ademais, deve-se ressaltar que das 97 operações levantadas, foram consideradas aptas ao estudo apenas 50 operações. As causas da eliminação das 47 operações restantes são: (i) opção de se trabalhar apenas com os

projetos caracterizados como “implantação”; (ii) indisponibilidade de informações confiáveis; e, (iii) prazo de maturação dos empreendimentos.³³

4.1. MODELOS DE RESPOSTA QUALITATIVA

Objetivando-se a construção de um modelo para previsão de inadimplência com base em variáveis qualitativas, determinou-se que a variável dependente (Y) corresponderá à situação em que se encontra o crédito, adimplente ou inadimplente. No primeiro caso receberá o valor “Zero”, enquanto no segundo receberá o valor “Um”. As variáveis explicativas (X’s) do modelo estão listadas na tabela 5.

Tabela 5. Variáveis Explicativas do Modelo de Resposta Qualitativa

Variável	Descrição
<i>POR</i>	Porte do Projeto
<i>TER</i>	Recursos de Terceiros
<i>TAX</i>	Taxa de Juros Anual Média Utilizada nos Contratos
<i>PRA</i>	Prazo de Médio de Pagamento
<i>FAT</i>	Faturamento Anual Previsto
<i>LUX</i>	Investimento Total / Número de Unidades Habitacionais

Deve-se ressaltar que: (i) o item “Recursos de Terceiros” engloba recursos federal e estadual, além dos incentivos oferecidos pelo estado; (ii) a taxa de juros considerada foi uma ponderação das taxas adotadas pelo repassadores de recursos federal e estadual, utilizando como fator de ponderação o montante repassado [excluindo os componentes variáveis da taxa, os quais eram na maior parte a Taxa Referencial e a Taxa de Juros de Longo Prazo]; e (iii) o prazo médio de pagamento inclui os períodos de carência e amortização.

³³ Alguns empreendimentos ainda estavam no começo da carência. Portanto, não havia dados sobre pagamentos das parcelas do financiamento.

A inexistência de variáveis oriundas de índices financeiros, retiradas dos balanços e/ou balancetes, baseia-se no fato de que todos os projetos considerados para a análise objetivavam a implantação das empresas (hotéis), não possuindo, portanto, registros contábeis.

Algumas variáveis, de caráter microeconômico, que podem ser importantes na explicação do fenômeno da inadimplência nesse segmento, foram descartadas durante o processo de análise preliminar dos dados. As variáveis descartadas e os motivos para o descarte são apresentados na tabela 6.

Tabela 6. Variáveis Descartadas

Variável	Descrição	Motivo do Descarte
<i>TIP</i>	Tipo de Administração	Todos os projetos tinham administração familiar
<i>ESC</i>	Escolaridade do Empresário	Informação Inexistente
<i>EXP</i>	Experiência do Empresário	Informação Inexistente

4.1.1. MODELO DE PROBABILIDADE LINEAR

O modelo de probabilidade linear utilizado na previsão de inadimplência para os financiamentos concedidos pelo BANDES ao segmento hoteleiro é:

$$Y_i = \beta_1 + \beta_2 POR + \beta_3 TER + \beta_4 TAX + \beta_5 PRA + \beta_6 FAT + \beta_7 LUX + u_i$$

onde:

Y_i pode assumir dois valores, ou valor 0 (adimplente) ou valor 1 (inadimplente);

β_i são os parâmetros a serem estimados;

POR, *TER*, *TAX*, *PRA*, *FAT* e *LUX* são as variáveis explicativas; e

u_i representa o termo estocástico.

Utilizando o estimador de Mínimos Quadrados Ordinários (MQO), são obtidas as seguintes estimativas para os parâmetros β_i (tabela 7):

Tabela 7. Estimativas dos Parâmetros do Modelo de Probabilidade Linear

Parâmetros	Coeficientes	Desvio Padrão	Estatística t	Prob
β_1	-0,13602945	0,26575403	-0,511862229	0.6114
β_2	-2,10302E-07	1,45252E-07	-1,447841274	0.1549
β_3	2,09951E-07	2,01368E-07	1,042621664	0.3030
β_4	6,014889757	2,233912993	2,692535374	0.0101
β_5	0,004212169	0,003134353	1,343871971	0.1860
β_6	1,47294E-07	1,81947E-07	0,809543631	0.4227
β_7	3,29951E-06	3,80074E-06	0,868121318	0.3901

Apesar dos resultados das estimativas obtidas com o MQO guardarem certas particularidades para o caso de variáveis qualitativas, os resultados obtidos são satisfatórios. Os testes de significância demonstraram que a variável *TAX* (*taxa de juros média*) apresenta efeito positivo sobre o fenômeno da inadimplência, ao nível de 1%, enquanto as variáveis *POR* (*porte do projeto*) e *PRA* (*prazo médio de pagamento*), se considerássemos o nível de 20%, também seriam significativas. As demais variáveis não foram significativas ao nível de 20% (teste bilateral).

A estatística F (2,867688) confirmou, ao nível de 5%, que pelo menos uma variável explicativa afeta a variável explicada (Y). O coeficiente de determinação (0,2857), conforme exposto anteriormente, não é de muita valia para modelos de resposta qualitativa, podendo, portanto, ser descartado da análise.

Ademais, por se tratar de um modelo de previsão, objetiva-se conhecer o potencial de classificação do mesmo. Para tanto, utilizando as observações disponíveis sobre o segmento de hotelaria e definindo-se como ponto de corte para a situação de inadimplência o valor de 0,60,³⁴ ou seja,

$$Y = 1 \quad \text{Se } Y_i \text{ previsto} \geq 0,60$$

$$Y = 0 \quad \text{Se } Y_i \text{ previsto} < 0,60$$

observa-se que o modelo classificou corretamente 80% das operações realizadas (tabela 8).

Tabela 8. Modelo de Probabilidade Linear: Observação x Previsão

Observação	Y observado	Y previsto	Status da Classificação	Observação	Y observado	Y previsto	Status da Classificação
1	1	0,67	correta	26	1	0,58	errada
2	1	0,51	errada	27	0	0,53	correta
3	1	0,90	correta	28	0	0,59	correta
4	1	0,57	errada	29	0	0,44	correta
5	1	0,80	correta	30	1	0,41	errada
6	1	0,92	correta	31	1	0,84	correta
7	0	0,41	correta	32	1	0,81	correta
8	0	0,55	correta	33	1	0,88	correta
9	1	1,34	correta	34	1	0,31	errada
10	1	0,98	correta	35	1	0,44	errada
11	1	0,61	correta	36	0	0,43	correta
12	1	0,75	correta	37	0	0,38	correta
13	1	1,00	correta	38	1	0,94	correta
14	1	1,23	correta	39	0	0,31	correta
15	1	0,99	correta	40	1	0,63	correta
16	1	1,09	correta	41	1	0,38	errada
17	0	0,39	correta	42	1	0,41	errada
18	1	0,67	correta	43	1	0,77	correta
19	1	0,72	correta	44	0	0,65	errada
20	1	0,92	correta	45	1	0,69	correta
21	1	0,75	correta	46	1	0,78	correta
22	0	0,57	correta	47	0	0,32	correta
23	0	0,82	errada	48	1	0,96	correta
24	0	0,53	correta	49	0	0,54	correta
25	0	0,33	correta	50	1	0,98	correta

Observa-se ainda, que do total de 10 (dez) classificações realizadas de forma equivocada, ao nível de corte determinado, 8 (oito) dizem respeito ao erro do “tipo 1”, ou seja, classificar um projeto como sendo de reduzido risco,

³⁴ O valor 0,60 foi definido com base no melhor resultado obtido pelos modelos lógite e próbite, qual seja,

quando na realidade não o é, e 2 (duas) referem-se ao erro do “tipo 2”, qual seja, classificar um projeto como sendo de elevado risco, quando na realidade o risco é baixo.

Deve-se ressaltar, novamente, que a possibilidade de obter probabilidades de inadimplência superiores a 100% ou inferiores a 0% é uma das características que inibem a utilização da probabilidade linear em modelos de previsão, sendo que neste caso específico, tal fato ocorreu em três situações.

4.1.2. MODELOS PRÓBITE E LÓGITE

Nas tabelas 9 e 10 encontram-se, respectivamente, as estimativas para os parâmetros dos modelos próbite e lógite.

Tabela 9. Estimativas dos Parâmetros do Modelo Próbite

Parâmetros	Coeficientes	Desvio Padrão	Estatística z	Prob.
β_1	-2,072817	1,019491	-2.033188	0.0420
β_2	-1,75E-06	7,86E-07	-2,222208	0.0263
β_3	2,68E-06	1,40E-06	1,914161	0.0556
β_4	19,38661	8,214457	2,360060	0.0183
β_5	0,014249	0,011950	1,192355	0.2331
β_6	5,06E-07	6,74E-07	0,750935	0.4527
β_7	8,74E-06	1,60E-05	0,546894	0.5845

Tabela 10. Estimativas dos Parâmetros do Modelo Lógico

Parâmetros	Coeficientes	Desvio Padrão	Estatística z	Prob.
β_1	-3,316508	1,699975	-1,950916	0.0511
β_2	-2,71E-06	1,24E-06	-2,179307	0.0293
β_3	4,11E-06	2,22E-06	1,854994	0.0636
β_4	31,01858	14,13515	2,194428	0.0282
β_5	0,022795	0,019784	1,152197	0.2492
β_6	7.91E-07	1,11E-06	0,711392	0.4768
β_7	1,52E-05	2,62E-05	0,579988	0.5619

E ambas as regressões, observa-se que as variáveis *POR* (*porte do projeto*) e *TAX* (*taxa de juros média*) são significativas a 5%, enquanto *TER* (*recursos de terceiros*) foi significativa a 10%.

Utilizando as observações disponíveis sobre o segmento de hotelaria e definindo-se como ponto de corte para a situação de inadimplência o valor de 0,60, ou seja,

$$Y = 1 \quad \text{Se } Y_i \text{ previsto} \geq 0,60$$

$$Y = 0 \quad \text{Se } Y_i \text{ previsto} < 0,60$$

percebe-se que ambos os modelos classificaram corretamente 82% das operações analisadas, ou seja, nas 50 observações utilizadas, os modelos previram de forma correta 41 resultados, significando um percentual de ganho na análise de 43,75%.³⁵

Observa-se, na tabela 11, que do total de 9 (nove) classificações realizadas de forma equivocada pelo modelo próbite, ao nível de corte

³⁵ Anteriormente se trabalhava com uma informação genérica de que 16 projetos, de um total de 50, tornaram-se inadimplentes.

determinado, 6 (seis) dizem respeito ao erro do “tipo 1”, ou seja, classificar um projeto como sendo de reduzido risco, quando na realidade não o é, e 3 (três) referem-se ao erro do “tipo 2”, qual seja, classificar um projeto como sendo de elevado risco, quando na realidade o risco é baixo.

Tabela 11. Modelo Próbite: Observação x Previsão

<i>Observação</i>	<i>Y observado</i>	<i>Y previsto</i>	<i>Status da Classificação</i>	<i>Observação</i>	<i>Y observado</i>	<i>Y previsto</i>	<i>Status da Classificação</i>
1	1	0,62	correta	26	1	0,61	correta
2	1	0,54	errada	27	0	0,53	correta
3	1	0,91	correta	28	0	0,62	errada
4	1	0,60	correta	29	0	0,36	correta
5	1	0,84	correta	30	1	0,36	errada
6	1	0,97	correta	31	1	0,97	correta
7	0	0,37	correta	32	1	0,84	correta
8	0	0,58	correta	33	1	0,87	correta
9	1	1,00	correta	34	1	0,28	errada
10	1	0,99	correta	35	1	0,48	errada
11	1	0,60	correta	36	0	0,28	correta
12	1	0,92	correta	37	0	0,12	correta
13	1	1,00	correta	38	1	0,95	correta
14	1	0,99	correta	39	0	0,27	correta
15	1	0,95	correta	40	1	0,72	correta
16	1	0,98	correta	41	1	0,36	errada
17	0	0,37	correta	42	1	0,40	errada
18	1	0,94	correta	43	1	0,83	correta
19	1	0,63	correta	44	0	0,71	errada
20	1	0,98	correta	45	1	0,74	correta
21	1	0,87	correta	46	1	0,88	correta
22	0	0,58	correta	47	0	0,28	correta
23	0	0,78	errada	48	1	0,96	correta
24	0	0,53	correta	49	0	0,57	correta
25	0	0,30	correta	50	1	0,92	correta

No tocante ao modelo lógite, tabela 12, constam as observações classificadas de forma equivocada, no total de 9, que coincidem com as classificações realizadas pelo modelo próbite. Sendo assim, 6 (seis) erros são do “tipo 1” e 3 (três) referem-se ao erro do “tipo 2”.

Tabela 12. Modelo Lógico: Observação x Previsão

Observação	Y observado	Y previsto	Status da Classificação	Observação	Y observado	Y previsto	Status da Classificação
1	1	0,63	correta	26	1	0,61	correta
2	1	0,54	errada	27	0	0,54	correta
3	1	0,90	correta	28	0	0,62	errada
4	1	0,60	correta	29	0	0,36	correta
5	1	0,83	correta	30	1	0,37	errada
6	1	0,95	correta	31	1	0,95	correta
7	0	0,37	correta	32	1	0,83	correta
8	0	0,58	correta	33	1	0,86	correta
9	1	1,00	correta	34	1	0,28	errada
10	1	0,98	correta	35	1	0,47	errada
11	1	0,61	correta	36	0	0,28	correta
12	1	0,90	correta	37	0	0,14	correta
13	1	0,99	correta	38	1	0,93	correta
14	1	0,98	correta	39	0	0,27	correta
15	1	0,94	correta	40	1	0,71	correta
16	1	0,96	correta	41	1	0,36	errada
17	0	0,37	correta	42	1	0,40	errada
18	1	0,92	correta	43	1	0,82	correta
19	1	0,64	correta	44	0	0,70	errada
20	1	0,97	correta	45	1	0,74	correta
21	1	0,86	correta	46	1	0,87	correta
22	0	0,59	correta	47	0	0,29	correta
23	0	0,78	errada	48	1	0,94	correta
24	0	0,53	correta	49	0	0,57	correta
25	0	0,30	correta	50	1	0,91	correta

Levando-se em consideração que ambos os modelos apresentam uma "tendência" maior ao erro "tipo 1", poder-se-ia deslocar o ponto de corte para cima, ou seja, auferindo maior rigor para a análise e reduzindo o risco do credor. Em compensação, aumentaria a probabilidade do erro do "tipo 2", elevando-se o risco do cliente, ou seja, algumas operações com histórico de inadimplência poderiam ser caracterizadas como inadimplentes.

A opção de utilizar apenas um nível de corte pode ser um tanto radical em certas situações. É possível que a utilização de faixas de escores (classes de riscos) seja mais recomendável na implantação de um modelo de previsão de inadimplência. Desse modo, trabalhar-se-ia, por exemplo, com três classes de risco: a primeira contemplaria propostas com escores elevados, as quais seriam aprovadas automaticamente; a segunda poderia ser caracterizada como

um região “nebulosa”, estariam propostas com necessidade de análise mais detalhada; e, finalmente, a terceira representaria propostas com baixos valores de escores, sendo por isso recusadas automaticamente.

4.2. MODELO BASEADO NO CONCEITO DE ENTROPIA

A entropia inicial nas operações de crédito realizadas pelo BANDES ao segmento de hotelaria, conforme eq. 27, é representada por:³⁶

$$Entropia = -\left(\frac{34}{50}\right)\log_2\left(\frac{34}{50}\right) - \left(\frac{16}{50}\right)\log_2\left(\frac{16}{50}\right)$$

Entropia inicial = 0,90438 bits

onde $\left(\frac{34}{50}\right)$ representa a frequência de ocorrência do fenômeno “inadimplência” e $\left(\frac{16}{50}\right)$ o seu complemento, ou seja, a não ocorrência da inadimplência.

Observa-se, através do resultado calculado, que a entropia inicial do sistema é elevada, revelando uma grande incerteza a respeito dos eventos a que estão sujeitas as operações de crédito (inadimplência ou adimplência).

Objetivando-se o decréscimo da entropia (ganho de informação), construiu-se um modelo, com base no método ID3, no qual são utilizadas 04 variáveis características dos empreendimentos financiados pelo BANDES, sendo uma variável com atributo não-numérico e as demais com atributos numéricos, conforme exposto na tabela 13.

³⁶ Considerou-se, para o cálculo da entropia do sistema, apenas as 50 operações de crédito analisadas neste estudo.

Tabela 13. Variáveis Explicativas do Modelo Utilizando o Conceito de Entropia

Variável	Descrição	Atributo
POR	Porte do Projeto	Numérico
LOC	Localização do Projeto	Não-numérico
TER	Endividamento Geral (Recursos de Terceiros)	Numérico
LUX	Investimento Total / Número de Unidades Habitacionais	Numérico

A classificação dos empreendimentos, necessária à formação dos conjuntos com características semelhantes, respeitou os seguintes critérios:³⁷

- i. referente ao tamanho: investimentos superiores a R\$ 1.000.000,00 foram considerados de “*grande porte*”; investimentos superiores a R\$ 300.000,00 e inferiores a R\$ 1.000.000,00 foram considerados de “*médio porte*”; e investimentos inferiores a R\$ 300.000,00 foram considerados de “*pequeno porte*”;
- ii. referente a localização: aos investimentos localizados na Região Serrana atribuiu-se o valor “1”; aos investimentos localizados no Litoral Sul o valor “2”; aos localizados no Litoral Norte o valor “3”; aos localizados na Grande Vitória o valor “4”; e para as demais regiões o valor “5”;
- iii. referente ao endividamento total: o endividamento foi considerado “*alto*” se a participação de capital de terceiros ultrapassou 60% do valor total do investimento; considerado “*médio*” se ultrapassou a 40% do total do investimento, mas foi inferior a 60%; e considerado “*baixo*” se ficou abaixo de 40% do total investido;

- iv. referente ao total investido por unidade habitacional: se o total investido por unidade habitacional foi superior a R\$ 40.000,00 considerou-se como um investimento *“alto”*; se superior a R\$ 20.000,00 e inferior a R\$ 40.000,00 considerou-se como um investimento *“médio”*; e abaixo de R\$ 20.000,00 como investimento *“baixo”*.

Após a definição das variáveis a serem utilizadas e da classificação dos empreendimentos, passa-se à construção das “árvores discriminatórias”, levando-se em consideração todas as possibilidades de permuta entre as referidas variáveis, ou seja, 24 possibilidades.³⁸

A aplicação do método de cálculo ID3 em cada uma das 24 possibilidades, conforme demonstrado no Apêndice 5, indica que ao se agrupar os empreendimentos em conjuntos com características semelhantes, obtém-se, no máximo, um ganho de informação de 0,6865 *bits*, ou seja, a entropia do sistema passa a ser, no mínimo, de 0,2179 *bits*. O exemplo a seguir (Árvore Discriminatória I) demonstra um caso específico em que o ganho de entropia com a ordenação das variáveis foi máximo.

³⁷ Definidas em conjunto com técnicos do BANDES.

³⁸ O número de permutações é igual ao fatorial de 4 (4!), mais especificamente, $4 \times 3 \times 2 \times 1$.

baixo. O valor da entropia nesse “ramo” é igual a

$$-\left(\frac{3}{7}\right)\log_2-\left(\frac{3}{7}\right)-\left(\frac{4}{7}\right)\log_2-\left(\frac{4}{7}\right)=0,98523 \text{ bits};$$

2. empreendimentos com endividamento *médio*, de *médio* porte, localizados na região 2 e com investimento por unidade habitacional *médio*. O valor da entropia nesse “ramo” é igual a

$$-\left(\frac{1}{2}\right)\log_2-\left(\frac{1}{2}\right)-\left(\frac{1}{2}\right)\log_2-\left(\frac{1}{2}\right)=1,00 \text{ bit},$$

3. empreendimentos com endividamento *médio*, de *pequeno* porte, localizados na região 2 e com investimento por unidade habitacional *baixo*. O valor da entropia nesse “ramo” é igual a

$$-\left(\frac{1}{2}\right)\log_2-\left(\frac{1}{2}\right)-\left(\frac{1}{2}\right)\log_2-\left(\frac{1}{2}\right)=1,00 \text{ bit},$$

A entropia para o conjunto “*endividamento alto – pequeno porte – localização 2*” é:

$$\left[\left(\frac{1}{8} \times 0\right) + \left(\frac{7}{8} \times 0,98523\right)\right] = 0,8621 \text{ bits}$$

A entropia para o conjunto “*endividamento alto – pequeno porte*” é

$$\left[\left(\frac{1}{17} \times 0\right) + \left(\frac{8}{17} \times 0,8621\right) + \left(\frac{5}{17} \times 0\right) + \left(\frac{2}{17} \times 0\right) + \left(\frac{1}{17} \times 0\right)\right] = 0,4057 \text{ bits}$$

A entropia para o conjunto “*endividamento alto*” é:

$$\left[\left(\frac{10}{30} \times 0\right) + \left(\frac{3}{30} \times 0\right) + \left(\frac{17}{30} \times 0,4057\right)\right] = 0,2299 \text{ bits}$$

Seguindo-se a mesma linha de raciocínio para *endividamento baixo* e *endividamento médio* encontra-se, respectivamente, 0 e 0,25 de entropia. Desta forma, tem-se a seguinte entropia para o sistema como um todo:

$$\left[\left(\frac{30}{50} \times 0,2299 \right) + \left(\frac{4}{50} \times 0 \right) + \left(\frac{16}{50} \times 0,25 \right) \right] = 0,2179$$

Assim, de uma entropia inicial de 0,90438 *bits* chega-se a uma entropia final de 0,2179 *bits*, ou seja, houve um ganho de informação da ordem de 0,6865 *bits*. A simples agregação dos empreendimentos, caracterizados por apenas 04 atributos, forneceu uma redução significativa na “desordem” do sistema.

A aplicabilidade do modelo baseado no conceito de entropia pode ser demonstrada na seguinte situação: um empresário sugere ao BANDES que o mesmo financie 70% dos investimentos necessários à implantação de um hotel com 20 unidades habitacionais na Região Serrana do Espírito Santo. O investimento total é de aproximadamente R\$ 1.200.000,00.

Ao se adaptar a proposta ao modelo de análise sugerido, ter-se-ia a seguinte situação: *endividamento alto – grande porte – localização 2 – investimento por unidade habitacional alto*. Com base no histórico de operações do BANDES, deve-se rejeitar a proposta, uma vez que todos os empreendimentos com essas características encontram-se inadimplentes com o banco.

Evidentemente que as propostas podem ser analisadas com base em qualquer uma das 24 possibilidades, devendo-se, apenas, atentar para a preferência por aquelas que apresentam maior ganho de informação.

5. Conclusões

As mudanças ocorridas na economia brasileira na década de 90, com o início efetivo da abertura comercial e estabilização econômica, foram responsáveis por uma série de transformações no sistema financeiro nacional. As instituições financeiras viram-se obrigadas, dentre outras coisas, a realizar uma reengenharia dos seus processos de análise de crédito, buscando maior precisão dos mesmos.

Neste contexto, a carteira de empréstimos das instituições financeiras ganhou importância frente aos demais ativos e seu gerenciamento eficaz é primordial para a manutenção e/ou ampliação da lucratividade. A inovação nos métodos de análise, utilizando-se, principalmente, de métodos estatísticos mais elaborados é fundamental para o alcance de tais objetivos.

No caso específico do BANDES, instituição voltada quase que exclusivamente para créditos de longo prazo, urge a necessidade de atualização dos processos de análise, principalmente se considerarmos também a elevada inadimplência observada em sua carteira de empréstimos e os resultados negativos registrados em seus balanços nos últimos anos.

Obviamente que a proposta metodológica e os modelos apresentados nesse trabalho não seriam capazes, por si só, de reduzir a zero a

inadimplência observada na carteira de empréstimos do BANDES ao segmento de hotelaria. Contudo podem auxiliar sobremaneira a redução desta inadimplência a níveis aceitáveis.

Deve-se também ressaltar que os modelos propostos apresentam algumas deficiências em sua elaboração, quais sejam: (i) na seleção inicial dos projetos considerou-se apenas aqueles que receberam apoio financeiro do BANDES, ficando de fora projetos rejeitados pelos analistas do banco;³⁹ (ii) não foram analisadas as garantias oferecidas, as quais podem influenciar bastante a decisão sobre a concessão de um crédito.

Quanto aos resultados apresentados pelos modelos, pode-se afirmar que ficaram dentro das expectativas, revelando ganhos significativos de informação. A síntese dos resultados pode ser enumerada de forma bem objetiva, como segue:

- 1) os modelos próbite e lógite apresentaram resultados mais eficientes em comparação ao modelo de probabilidade linear;
- 2) o modelo baseado no conceito de entropia apresentou bons resultados, mas sua aplicabilidade é reduzida, face a dificuldades em alterar/refazer as "árvores discriminatórias";
- 3) as variáveis escolhidas conseguiram explicar de maneira satisfatória o problema da inadimplência, com destaque para a variável "taxa de juros";

Poder-se-ia também realizar uma análise mais simples do fenômeno inadimplência no segmento de hotelaria apoiado pelo BANDES, utilizando, por exemplo, uma análise via estatística descritiva das variáveis explicativas, conforme demonstrado na tabela 14.

Tabela 14. Análise Estatística Descritiva das Variáveis Explicativas

Variável	Média		Total
	Adimplentes	Inadimplentes	
<i>POR</i>	R\$ 672.993	R\$ 1.446.157	R\$ 1.198.744
<i>TER</i>	R\$ 316.314	R\$ 957.878	R\$ 752.577
<i>TAX</i>	5.50%	8.32%	7,42%
<i>PRA</i>	58 meses	67 meses	64 meses
<i>FAT</i>	R\$ 441.160	R\$ 633.291	R\$ 571.809
<i>LUX</i>	R\$ 17.187	R\$ 40.607	R\$ 33.113

Observa-se claramente, através das médias, que os projetos inadimplentes têm como características marcantes:

- 1) valor de implantação [POR] elevado;
- 2) percentual de capitais de terceiros [TER] elevado;;
- 3) taxa de juros média [TAX] elevada;
- 4) investimento por unidade habitacional [LUX] elevado;

Por outro lado, observa-se que os projetos adimplentes têm como características marcantes:

- 1) valor de implantação [POR] reduzido;
- 2) percentual de capitais de terceiros [TER] em níveis menores;;
- 3) taxa de juros média [TAX] reduzida;
- 4) investimento por unidade habitacional [LUX] reduzido;

Retornando aos modelos apresentados nesse estudo, deve-se ainda tecer algumas considerações. A primeira diz respeito a não necessidade da utilização dos modelos lógite e próbite de forma conjunta, uma vez que apresentam resultados similares. Greene (1993), faz menção ao uso exclusivo de somente um desses modelos.

³⁹ Não haviam registros e/ou arquivos sobre os projetos rejeitados pelo BANDES.

Relembrando a referência feita a Amemiya (1981), na qual o mesmo afirma que pode-se obter valores aproximados dos coeficientes da distribuição logística multiplicando os coeficientes do modelo próbite pela constante 1,60 e que os coeficientes do modelo de probabilidade linear podem ser obtidos multiplicando-se os coeficientes do modelo lógite por 0,25 (exceto para a constante, que seria obtida pela multiplicação do respectivo coeficiente por 0,25, devendo-se ainda ao valor final somar 0,5), chega-se também a conclusão, que ao menos para o caso apresentado nesse estudo, poder-se-ia adotar o modelo Próbite para em seguida obter os coeficientes do modelo Lógite [e vice-versa]. Contudo, para o modelo de probabilidade linear a afirmação não foi comprovada.

Tabela 15. Coeficientes de Regressão

Parâmetros	MPL	Próbite	Lógite
β_1	-0,13602945	-2,072817	-3,316508
β_2	-2,10302E-07	-1,75E-06	-2,71E-06
β_3	2,09951E-07	2,68E-06	4,11E-06
β_4	6,014889757	19,38661	31,01858
β_5	0,004212169	0,014249	0,022795
β_6	1,47294E-07	5,06E-07	7,91E-07
β_7	3,29951E-06	8,74E-06	1,52E-05

Finalizando, deve-se considerar que os resultados obtidos pelos modelos propostos foram satisfatórios e podem ajudar sobremaneira na identificação e quantificação das variáveis explicativas do fenômeno da inadimplência nos empréstimos realizados pelo BANDES ao segmento hoteleiro do Espírito Santo.

Deve-se ainda acrescentar que pela eficiência nas respostas e pela facilidade de utilização e adaptação, o modelo Lógite seria o mais indicado para

análises similares à apresentada nesse estudo. Devendo também ser considerado pelo BANDES a identificação e quantificação das variáveis explicativas para os demais segmentos com problemas de inadimplência em sua carteira de empréstimos.

Referências Bibliográficas

- ALTMAN, Edward I. Financial Ratios, Discriminants Analysis and the Prediction of Corporate Bankruptcy. **Journal of Finance**, Sep/1968.
- ALTMAN, E. I; HALDEMAN, R. G.; NARAYANAN, P. Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations. **Journal of Banking and Finance**, 1, 29-54, 1977.
- AMEMIYA, Takeshi. Qualitative Response Models: A Survey. **Journal of Economic Literature**, XIX, p. 1483-1536, 1981.
- _____. **Advanced Econometrics**. Cambridge: Harvard University Press, 1985.
- ANGBAZO, Lazarus. Commercial Bank Net Interest Margins, Default Risk, Interest-Rate Risk, and Off-Balance Sheet Banking. **Journal of Banking and Finance**, 21, 55-87, 1997.
- ARAÚJO, Uilson Melo. **Assimetria de Informação no Crédito Rural: Aspectos Teóricos e um Modelo para Classificação do Risco dos Créditos Concedidos a Cooperativas Agropecuárias**. Piracicaba: ESALQ/USP, Tese de Doutorado, 1996.
- BANCO CENTRAL DO BRASIL. **Boletim do Banco Central do Brasil**. Brasília, Abril, 1997.
- _____. **O Banco Central e os Dois Anos do Real**. Brasília, 1996.
- BARROS, J. R. M.; ALMEIDA JÚNIOR, Mansueto F. **A Reestruturação do Sistema Financeiro no Brasil**. (mimeo). Brasília, Agosto, 1996.

- BOYES, W. J.; HOFFMAN, D. L.; LOW, S. An Econometric Analysis of the Bank Credit Scoring Problem. **Journal of Econometrics**, 40, 3-14, 1989.
- CAPON, Noel. Credit Scoring Systems: A Critical Analysis. **Journal of Marketing**, 46, 82-91, 1982.
- DeVANY, A. S. Comment on Modeling in Banking Firm: A Survey. **Journal of Money, Credit and Banking**, 16-4, 603-609, 1984.
- FONS, Jerome S. Using Default Rates to Model the Term Structure of Credit Risk. **Financial Analysis Journal**, 25-32, Sep/Oct, 1994.
- GOLDEN, Sam; WALKER, H. M. The Ten Commandments of Commercial Credit: The Cs of Good and Bad Loans. **The Journal of Lending & Credit Risk Management**, 14-20, May, 1996.
- GRABLOWSKY, B. J.; TALLEY, W. K. Probit and Discriminant Functions for Classifying Credit Applicants: A Comparasion. **Journal of Economics and Business**, 33, 254-261, 1981.
- GREENE, William H. **Econometric Analysis**. New Jersey: Prentice-Hall, Second Edition, 1993.
- GREENE, William H. **LINDEP User's Manual: Version 7.0**. New York: Econometric Software, 1995.
- GUJARATI, Damodar N. **Basic Econometrics**. New York: MacGraw-Hill, Third Edition, 1995.
- KANITZ, Sthephen Charles. **Como Prever Falências**. São Paulo: McGraw-Hill, 1978.
- KEYNES, J. M. **A Teoria Geral do Emprego, do Juro e da Moeda**. São Paulo: Editora Atlas, 1990.
- KOH, Hian Chye; TAN, Kim Wah. Construction of a credit scoring model to assess the default risk of credit card applicants. **International Journal of Management**, 11-3, p. 746-751, Sep/1994.
- LAUDEMAN, Mark. Risk-Rating Systems Bring Consistency to Commercial Lending. **Commercial Lending Review**, 1V.

- LAWRENCE, Edward C.; NASSER Arshadi. A Multinomial Logit Analysis of Problem Loan Resolution Choices in Banking. **Journal of Money, Credit and Banking**, 27-1, Feb, 1995.
- LENTINO, J. V.; RIZZI, J. V. Credit Analysis for Highly Leveraged Credits: Déjà Vu All Over Again. **Commercial Lending Review**, 11, 21-28, Winter 1995/1996.
- LEONARD, K. J.; BANKS, W. J. Automating the Credit Decision Process. **Journal of Retail Banking**, XVI -1, 39-44, 1994.
- LIMA, Ricardo C. Modelos de Respostas Binárias: Especificação, Estimação e Inferência. **Agricultura em São Paulo**, 43 (2): 19-25, 1996.
- MADDALA, G. **Limited Dependent and Qualitative Variables in Econometrics**. New York: Cambridge University Press, 1983.
- McLEOD, R. W.; MALHOTRA, D. K.; MALHOTRA, R. Predicting Credit Risk: A Neural Network Approach. **Journal of Retail Banking**, XV-3, 37-40, 1993.
- ORGLER, Yair E. A Credit Scoring Model for Commercial Loans. **Journal of Money, Credit and Banking**, 2, 435-445, Nov/1970.
- PAO, Yoh-Han. *Adaptive Pattern Recognition and Neural Networks*. New York: Addison-Wesley Publishing Company, 1989.
- PESSOA, A. S. M. **Custos Bancários de Transação do Crédito Rural no Brasil**. Piracicaba: DESR/ESALQ/USP. Dissertação de Mestrado, 1996.
- PINDER, Jonathan P. Decision analysis using multinomial logit models: Mortgage portfolio valuation. **Journal of Economics & Business**, 48-1, p. 67-77, Feb/1996.
- SANTOMERO, Anthony M. Modeling the Banking Firm: A Survey. **Journal of Money, Credit and Banking**, 16-4, 576-602, 1984.
- SECURATO, J. R. **Decisões Financeiras em Condições de Risco**. São Paulo: Editora Atlas, 1996.

- SILVA, José P. **Gestão e Análise de Risco de Crédito**. São Paulo: Editora Atlas, 1997.
- SMITH, L Douglas; SANCHEZ, Susan M; LAWRENCE, Edward C. A comprehensive model for managing credit risk on home mortgage portfolios. **Decision Sciences**, 27-2, p. 291-317, Spring 1996.
- WU, Hsiu-Kwang. Bank Examiner Criticisms, Bank Loan Defaults, and Bank Loan Quality. **Journal of Finance**, XXIV, 697-705, Sep/1969.

APÊNDICE 1: Dados

POR	TER	TAX	PRZ	FAT	LUX
543.263,34	190.141,58	8,00%	60,00	359.639,04	27.163,17
91.331,34	68.498,50	3,50%	92,00	91.160,03	11.416,42
282.659,90	179.535,00	8,00%	117,00	192.222,14	16.627,05
29.000,00	18.817,30	6,50%	72,00	52.850,00	1.812,50
217.360,00	120.000,00	8,00%	96,00	138.761,00	16.720,00
2.038.248,25	1.341.540,00	11,50%	60,00	698.021,16	48.529,72
187.615,91	81.949,00	4,00%	60,00	278.907,57	9.380,80
190.000,00	120.391,47	8,00%	27,00	600.000,00	4.750,00
5.472.528,00	5.284.516,80	11,50%	50,00	1.614.582,34	114.011,00
2.349.365,05	1.699.766,47	11,50%	57,00	459.134,80	78.312,17
832.635,14	416.317,57	8,00%	36,00	724.532,24	27.754,50
4.115.932,95	2.565.173,46	6,50%	72,00	981.720,94	114.331,47
7.021.496,13	4.645.872,81	8,50%	72,00	1.212.580,65	195.041,56
268.525,86	173.370,00	12,00%	120,00	189.666,78	38.360,84
5.057.788,85	2.819.717,27	11,50%	60,00	1.580.370,00	126.444,72
669.208,20	404.530,00	12,00%	96,00	318.941,85	33.460,41
75.460,00	50.000,00	4,00%	60,00	62.207,68	7.546,00
4.756.695,03	3.091.075,83	6,50%	72,00	1.519.272,24	72.071,14
3.759.649,00	1.822.911,60	11,50%	48,00	1.746.673,30	34.811,56
881.074,74	857.353,58	11,50%	48,00	951.600,00	8.158,10
1.622.627,65	1.032.581,23	12,00%	56,00	44.179,21	15.453,60
346.328,54	173.370,00	3,50%	87,00	626.559,18	21.645,53
346.500,00	60.000,00	11,50%	51,00	315.047,69	19.250,00
208.021,22	107.721,00	6,00%	60,00	208.021,22	11.556,73
77.588,00	55.420,00	3,50%	56,00	98.758,44	2.216,80
73.470,00	58.776,00	8,00%	48,00	53.632,28	7.347,00
325.454,60	163.898,00	3,50%	97,00	187.312,00	14.793,39
295.492,26	168.112,26	8,00%	48,00	169.840,00	13.431,47
501.861,52	184.279,60	4,50%	45,00	849.467,17	16.728,72
595.000,00	290.789,00	3,50%	51,00	722.541,89	23.800,00
1.781.879,08	1.335.196,77	8,00%	24,00	2.335.179,60	42.425,69
364.575,74	179.535,00	8,00%	96,00	195.453,77	22.785,98
425.444,66	192.660,00	8,00%	93,00	265.811,72	47.271,63
30.045,87	23.084,60	4,00%	45,00	56.452,34	3.004,59
271.902,17	203.926,63	3,50%	56,00	802.401,76	6.323,31
6.555.000,00	3.356.160,00	8,00%	56,00	1.798.562,30	77.117,65
890.493,60	83.783,00	4,00%	57,00	937.895,62	19.788,75
814.123,20	549.366,05	11,50%	56,00	947.656,43	20.353,08
55.047,20	26.465,00	4,00%	45,00	37.051,00	5.504,72
160.384,04	160.384,04	8,00%	56,00	185.472,00	5.346,13
90.318,07	59.845,00	4,00%	57,00	69.898,96	9.031,81
252.775,20	151.664,19	4,50%	57,00	200.906,17	7.222,15
277.764,82	179.535,00	8,00%	84,00	274.586,98	15.431,38
245.364,50	155.597,00	8,00%	56,00	235.789,30	17.526,04
263.318,00	167.566,00	8,00%	58,00	181.928,80	29.257,56
1.738.530,00	1.119.920,66	8,00%	42,00	1.123.200,00	66.866,54
260.053,12	149.321,52	4,00%	39,00	221.996,00	13.002,66
1.116.567,10	725.768,62	11,50%	56,00	892.970,12	37.218,90
207.610,00	124.565,54	3,50%	84,00	431.148,13	20.761,00
903.844,48	438.120,00	8,00%	118,00	347.925,70	56.490,28

APÊNDICE 2: Estimativas das regressões do modelo de probabilidade linear

Dependent Variable: SIT
 Method: Least Squares
 Date: 01/11/00 Time: 00:29
 Sample: 1 50
 Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.136029	0.265754	-0.511862	0.6114
POR	-2.10E-07	1.45E-07	-1.447841	0.1549
TER	2.10E-07	2.01E-07	1.042622	0.3030
TAX	6.014890	2.233913	2.692535	0.0101
PRZ	0.004212	0.003134	1.343872	0.1860
FAT	1.47E-07	1.82E-07	0.809544	0.4227
LUX	3.30E-06	3.80E-06	0.868121	0.3901
R-squared	0.285787	Mean dependent var		0.680000
Adjusted R-squared	0.186129	S.D. dependent var		0.471212
S.E. of regression	0.425103	Akaike info criterion		1.256206
Sum squared resid	7.770638	Schwarz criterion		1.523889
Log likelihood	-24.40516	F-statistic		2.867688
Durbin-Watson stat	1.349484	Prob(F-statistic)		0.019366

APÊNDICE 3: Estimativas das regressões do modelo próbite

Dependent Variable: SIT
 Method: ML - Binary Probit
 Date: 01/11/00 Time: 00:23
 Sample: 1 50

Included observations: 50

Convergence achieved after 1 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-2.072817	1.019491	-2.033188	0.0420
POR	-1.75E-06	7.86E-07	-2.222208	0.0263
TER	2.68E-06	1.40E-06	1.914161	0.0556
TAX	19.38661	8.214457	2.360060	0.0183
PRZ	0.014249	0.011950	1.192355	0.2331
FAT	5.06E-07	6.74E-07	0.750935	0.4527
LUX	8.74E-06	1.60E-05	0.546894	0.5845
Mean dependent var	0.680000	S.D. dependent var	0.471212	
S.E. of regression	0.404531	Akaike info criterion	1.113171	
Sum squared resid	7.036762	Schwarz criterion	1.380854	
Log likelihood	-20.82927	Hannan-Quinn criter.	1.215106	
Restr. log likelihood	-31.34347	Avg. log likelihood	-0.416585	
LR statistic (6 df)	21.02840	McFadden R-squared	0.335451	
Probability(LR stat)	0.001813			
Obs with Dep=0	16	Total obs	50	
Obs with Dep=1	34			

Dependent Variable: SIT

Method: ML - Binary Probit

Date: 01/11/00 Time: 00:23

Sample: 1 50

Included observations: 50

Prediction Evaluation (success cutoff C = 0.6)

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	13	6	19	0	0	0
P(Dep=1)>C	3	28	31	16	34	50
Total	16	34	50	16	34	50
Correct	13	28	41	0	34	34
% Correct	81.25	82.35	82.00	0.00	100.00	68.00
% Incorrect	18.75	17.65	18.00	100.00	0.00	32.00
Total Gain*	81.25	-17.65	14.00			
Percent Gain**	81.25	NA	43.75			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	8.76	7.50	16.26	5.12	10.88	16.00
E(# of Dep=1)	7.24	26.50	33.74	10.88	23.12	34.00
Total	16.00	34.00	50.00	16.00	34.00	50.00
Correct	8.76	26.50	35.25	5.12	23.12	28.24
% Correct	54.72	77.93	70.51	32.00	68.00	56.48
% Incorrect	45.28	22.07	29.49	68.00	32.00	43.52
Total Gain*	22.72	9.93	14.03			
Percent Gain**	33.42	31.04	32.23			

*Change in "% Correct" from default (constant probability) specification

**Percent of incorrect (default) prediction corrected by equation

APÊNDICE 4: Estimativas das regressões do modelo lógico

Dependent Variable: SIT
 Method: ML - Binary Logit
 Date: 01/11/00 Time: 00:14
 Sample: 1 50

Included observations: 50

Convergence achieved after 1 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-3.316508	1.699975	-1.950916	0.0511
POR	-2.71E-06	1.24E-06	-2.179307	0.0293
TER	4.11E-06	2.22E-06	1.854994	0.0636
TAX	31.01858	14.13515	2.194428	0.0282
PRZ	0.022795	0.019784	1.152197	0.2492
FAT	7.91E-07	1.11E-06	0.711392	0.4768
LUX	1.52E-05	2.62E-05	0.579988	0.5619
Mean dependent var	0.680000	S.D. dependent var	0.471212	
S.E. of regression	0.406077	Akaike info criterion	1.126713	
Sum squared resid	7.090643	Schwarz criterion	1.394396	
Log likelihood	-21.16782	Hannan-Quinn criter.	1.228648	
Restr. log likelihood	-31.34347	Avg. log likelihood	-0.423356	
LR statistic (6 df)	20.35130	McFadden R-squared	0.324650	
Probability(LR stat)	0.002397			
Obs with Dep=0	16	Total obs	50	
Obs with Dep=1	34			

Dependent Variable: SIT
 Method: ML - Binary Logit
 Date: 01/11/00 Time: 00:14
 Sample: 1 50

Included observations: 50

Prediction Evaluation (success cutoff C = 0.6)

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)<=C	13	6	19	0	0	0
P(Dep=1)>C	3	28	31	16	34	50
Total	16	34	50	16	34	50
Correct	13	28	41	0	34	34
% Correct	81.25	82.35	82.00	0.00	100.00	68.00
% Incorrect	18.75	17.65	18.00	100.00	0.00	32.00
Total Gain*	81.25	-17.65	14.00			
Percent Gain**	81.25	NA	43.75			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	8.70	7.75	16.45	5.12	10.88	16.00
E(# of Dep=1)	7.30	26.25	33.55	10.88	23.12	34.00
Total	16.00	34.00	50.00	16.00	34.00	50.00
Correct	8.70	26.25	34.95	5.12	23.12	28.24
% Correct	54.37	77.20	69.89	32.00	68.00	56.48
% Incorrect	45.63	22.80	30.11	68.00	32.00	43.52
Total Gain*	22.37	9.20	13.41			
Percent Gain**	32.90	28.74	30.82			

*Change in "% Correct" from default (constant probability) specification

**Percent of incorrect (default) prediction corrected by equation

APÊNDICE 5: "Árvores Discriminatórias" do modelo baseado no conceito de entropia

Árvore Discriminatória I

Nível 01: Grau de Endividamento do Projeto
Nível 02: Porte do Projeto
Nível 03: Localização do Projeto
Nível 04: R\$ por Unidade Habitacional

Ganho de Entropia: 0.6865

Árvore Discriminatória II

Nível 01: Grau de Endividamento do Projeto

Nível 02: Porte do Projeto

Nível 03: R\$ por Unidade Habitacional

Nível 04: Localização do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória III

Nível 01: Porte do Projeto
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: R\$ por Unidade Habitacional
 Nível 04: Localização do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória IV

Nível 01: Porte do Projeto
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: Localização do Projeto
 Nível 04: R\$ por Unidade Habitacional

Ganho de Entropia: 0.6865

Árvore Discriminatória V

Nível 01: Porte do Projeto
 Nível 02: Localização do Projeto
 Nível 03: Grau de Endividamento do Projeto
 Nível 04: R\$ por Unidade Habitacional

Ganho de Entropia: 0.6865

Árvore Discriminatória VI

Nível 01: Porte do Projeto
 Nível 02: Localização do Projeto
 Nível 03: RS por Unidade Habitacional
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6856

Árvore Discriminatória VII

Nível 01: Porte do Projeto
 Nível 02: R\$ por Unidade Habitacional
 Nível 03: Grau de Endividamento do Projeto
 Nível 04: Localização do Projeto

Ganho de Entropia: 0.6856

Árvore Discriminatória VIII

Nível 01: Porte do Projeto
 Nível 02: R\$ por Unidade Habitacional
 Nível 03: Localização do Projeto
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória X

Nível 01: Grau de Endividamento do Projeto

Nível 02: Localização do Projeto

Nível 03: Porte do Projeto

Nível 04: RS por Unidade Habitacional

Ganho de Entropia: 0.6853

Árvore Discriminatória XI

Nível 01: Grau de Endividamento do Projeto

Nível 02: R\$ por Unidade Habitacional

Nível 03: Porte do Projeto

Nível 04: Localização do Projeto

Ganho de Entropia: 0.6688

Árvore Discriminatória XII

Nível 01: Grau de Endividamento do Projeto

Nível 02: RS por Unidade Habitacional

Nível 03: Localização do Projeto

Nível 04: Porte do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XIII

Nível 01: RS por Unidade Habitacional
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: Localização do Projeto
 Nível 04: Porte do Projeto

Ganho de Entropia: 0.6856

Árvore Discriminatória XIV

Nível 01: R\$ por Unidade Habitacional
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: Porte do Projeto
 Nível 04: Localização do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XV

Nível 01: RS por Unidade Habitacional
 Nível 02: Localização do Projeto
 Nível 03: Grau de Endividamento do Projeto
 Nível 04: Porte do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XVI

Nível 01: RS por Unidade Habitacional
 Nível 02: Localização do Projeto
 Nível 03: Porte do Projeto
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XVII

Nível 01: RS por Unidade Habitacional
 Nível 02: Porte do Projeto
 Nível 03: Localização do Projeto
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XVIII

Nível 01: RS por Unidade Habitacional
 Nível 02: Porte do Projeto
 Nível 03: Localização do Projeto
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XIX

Nível 01: Localização do Projeto
 Nível 02: Porte do Projeto
 Nível 03: RS por Unidade Habitacional
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6865

Árvore Discriminatória XX

Nível 01: Localização do Projeto
 Nível 02: Porte do Projeto
 Nível 03: Grau de Endividamento do Projeto
 Nível 04: RS por Unidade Habitacional

Ganho de Entropia: 0.6865

Árvore Discriminatória XXI

Nível 01: Localização do Projeto
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: Porte do Projeto
 Nível 04: RS por Unidade Habitacional

Ganho de Entropia: 0.6865

Árvore Discriminatória XXII

Nível 01: Localização do Projeto
 Nível 02: Grau de Endividamento do Projeto
 Nível 03: RS por Unidade Habitacional
 Nível 04: Porte do Projeto

Ganho de Entropia: 0.6856

Árvore Discriminatória XXIII

Nível 01: Localização do Projeto
 Nível 02: RS por Unidade Habitacional
 Nível 03: Grau de Endividamento do Projeto
 Nível 04: Porte do Projeto

Ganho de Entropia: 0.6856

Árvore Discriminatória XXIV

Nível 01: Localização do Projeto
 Nível 02: RS por Unidade Habitacional
 Nível 03: Porte do Projeto
 Nível 04: Grau de Endividamento do Projeto

Ganho de Entropia: 0.6855