



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

TRABALHO DE PROJETO

CREDIT SCORING: UMA ANÁLISE ECONOMETRICA

ANTÓNIO FRANCISCO DE MELO FERNANDES

OUTUBRO - 2017



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO EM
MÉTODOS QUANTITATIVOS PARA A DECISÃO
ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO
TRABALHO DE PROJETO

CREDIT SCORING: UMA ANÁLISE ECONOMETRICA

ANTÓNIO FRANCISCO DE MELO FERNANDES

ORIENTAÇÃO:

PROFESSORA DOUTORA ISABEL MARIA DIAS PROENÇA

PROFESSORA DOUTORA MARIA CÂNDIDA VERGUEIRO

MONTEIRO CIDADE MOURÃO

OUTUBRO - 2017

Agradecimentos

Primeiramente devo agradecer à Deus pela vida que Ele me concedeu.

À Direção da Universidade Técnica de Angola pela bolsa de estudo; à Direção do ISEG por este mestrado.

Às minhas orientadoras professora Doutora Isabel Proença e professora Doutora Maria Cândida Mourão pelos conhecimentos que me transmitiram e pela disponibilidade em esclarecerem as minhas dúvidas.

Aos meus pais Amaral Manuel Fernandes e Clementina António Francisco, pelos princípios e ensinamentos que guiaram minha educação e a família em geral.

À minha querida esposa Madalena de Melo Fernandes pelo carinho, compreensão, apoio e paciência durante minha ausência.

Aos colegas e amigos da bolsa que me acompanharam nesta jornada, em especial à Felicidade Quivuna.

Aos colegas do curso pelas alegrias e sufocos que passamos juntos, unidos incentivando e apoiando uns aos outros.

Lista de siglas e abreviatura

AD	Ativos Disponíveis
AUTO	Auto
CASA	Casa
CRED	Credibilidade
DC	Duração do Crédito
DESEMP	Desempregado
E/CP/QSFP	Executivo /por conta própria /quadro superior da função pública
EDUC	Educação
EMPM4	Empregado a 4 anos ou mais
EPM	Efeito parcial médio
FER	Férias
FPCA	Falhas no pagamento do crédito anterior
FPEQS	Funcionário público com exceção de quadro superior
IDADE	Idade
LD	Outros créditos na loja de departamento
LR	Likelihood ratio
MC	Montante do crédito
MCEOB	Mais créditos em execução em outros bancos
NCANB	Número de créditos anteriores neste banco (incluindo o que está em execução)
NCM6	Número de créditos maior ou igual a 6
ND	Número de dependentes
NDM3	Número de dependentes maior ou igual a 3
NEGO	Negócios
OB	Outros créditos em outros bancos
OC	Objetivo do crédito
OCE	Outros créditos em execução
OCUP	Ocupação
PCT	Posse de casa ou terra
PRESTM20 A 25	Prestação do pagamento do crédito maior ou igual a 20%
PRESTM25 A 35	Prestação do pagamento do crédito maior ou igual a 25% e inferior a 35%
PRESTM35	Prestação do pagamento do crédito maior ou igual a 35%
SMC	Sem movimento na conta
SPCE	Situação perante créditos existentes
SSC	Saldo de conta corrente
SSD	Sem saldo na conta ou débito
TAXA	Taxa
TE	Tempo de emprego

Resumo

Com intenção de melhorar os serviços de análise e gestão de crédito, as instituições financeiras desenvolveram o modelo *credit scoring*. Este modelo é utilizado por estas instituições para previsão do risco de crédito no processo da tomada de decisão de concessão de crédito. O objetivo deste trabalho, é desenvolver um modelo de *credit scoring* a partir de uma amostra de 1000 solicitantes de créditos extraídos da carteira de crédito de um banco alemão. Para tal, estimou-se um modelo *probit*, considerando-se 25 variáveis independentes quantitativas e qualitativas que influenciam a probabilidade do crédito ser aprovado ou não. Os resultados deste estudo mostram que o modelo de *credit scoring* se apresenta adequado no ajustamento aos dados, obtendo uma classificação correta para cerca de 77% dos clientes. Contudo, os resultados encontrados fornecem informações importantes para auxílio no processo de tomada de decisões de concessão de crédito e gerenciamento do crédito bancário, podendo assim contribuir para a redução do número de clientes inadimplentes e dos respetivos custos.

Palavras-chave: Crédito, Risco de Crédito, *Credit Scoring*, *Probit*.

Abstract

In order to improve credit analysis and management services, financial institutions have developed the credit scoring model. This model is used by these institutions to predict credit risk in the process of making a credit granting decision. The objective of this work is to develop a credit scoring model from a sample of 1000 credit claimants extracted from the credit portfolio of a German bank. For this, a probit model was estimated, considering 25 independent quantitative and qualitative variables that influence the probability of credit being approved or not. The results of this study show that the credit scoring model is adequate in the adjustment to the data, obtaining a correct classification for about 77% of the clients. However, the results found provide important information to aid in the decision-making process of credit granting and bank credit management, thus contributing to the reduction of overdue customers and their costs

Key words: Credit, Credit risk, Credit Scoring, Probit.

Índice

Agradecimentos	i
Lista de siglas e abreviatura	ii
Resumo	iii
Abstract.....	iv
Índice de tabelas	vii
Capítulo 1 Introdução	1
1.1 Motivação e Objetivos do Trabalho.....	2
1.2 Estrutura do Relatório	3
Capítulo 2 Enquadramento Teórico.....	4
2.1 O Conceito de Crédito	4
2.2. O Risco de Crédito.....	5
2.3 Causas do Risco de Crédito	6
2.4 Análise de Crédito	7
2.4.1 Análise Subjetiva do Crédito.....	9
2.4.2 Análise Objetiva de Crédito	11
2.4.3 Credit Scoring	11
2.5 Técnicas Estatísticas aplicadas em Modelos de <i>Credit Scoring</i>	14
2.6 O Modelo <i>Probit</i>	15
Capítulo 3 Aplicação Empírica	16
3.1 Apresentação da Amostra e das Variáveis.....	16
3.2 Estimação.....	23
3.3 Testes de Significância Conjunta no Modelo <i>Probit</i>	25
3.4 Interpretação dos Resultados	27
3.5 Estimação dos Efeitos Parciais	28

3.6 Análise do Poder Preditivo do Modelo	29
Capítulo 4 Conclusão	32
Referências Bibliográficas.....	34
Anexo 1 Resumo das variáveis da base de dados.....	36
Anexo 2 Estatística descritiva	39
Anexo 3 Correlação da variável dependente com outras variáveis	40
Anexo 4 Teste de significância.....	41
Anexo 5 Estimação do poder preditivo	43

Índice de tabelas

Tabela 1 Definição das Variáveis	20
Tabela 2 Modelos estimados	24
Tabela 3 Teste de significância	26
Tabela 4 Interpretação dos sinais.....	27
Tabela 5 Efeitos parciais médios	29
Tabela 6 Poder preditivo do modelo 2.....	30
Tabela 7 Frequência relativa das variáveis categóricas segundo o tipo de crédito	38
Tabela 8 Estatística descritiva	39
Tabela 9 Correlação.....	40
Tabela 10 Teste de significância conjunta de fer e nego	41
Tabela 11 Teste de significância conjunta de ob e ld	41
Tabela 12 Teste de significância conjunta de fpeqs e qsfp	42
Tabela 13 Teste de significância conjunta de todas não significativas	42
Tabela 14 Estimação do poder preditivo	43

Capítulo 1 Introdução

Grandes e pequenas empresas tomam empréstimos para satisfação de diversas necessidades, tais como compras de matérias-primas, expansão do negócio, expansão da capacidade produtiva, etc. Tal como as empresas, as pessoas também tomam empréstimos para satisfação das suas necessidades, assim como compra de casa, carros, férias, etc. Com a disponibilidade de empréstimos para todos os segmentos da sociedade, a economia é catapultada e o crédito é visto como uma ferramenta poderosa para o progresso económico. Porém, existe sempre um risco de crédito quando se toma um produto ou serviço sem fazer o pagamento imediato por este. Segundo (Bastardo, 2011), “a maioria das crises financeiras ocorridas nas últimas décadas, resultam quase sempre de problemas de crédito”. De facto, percebemos que a exposição ao risco de crédito tem aumentado gradualmente com as repentinas mudanças económicas, políticas e tecnológicas que vêm surgindo no mundo. Essas mudanças alteram, de certa forma, o processo de gestão do risco de crédito nas instituições financeiras. Tendo em conta que o “risco de crédito é considerado o mais importante de todos os riscos bancários” (Bessis, 2010), as instituições financeiras procuram constantemente melhorar o processo de gestão do risco de crédito. Para tal, buscam constantemente aperfeiçoar ou desenvolver novas técnicas para gestão do risco de crédito. De entre estas técnicas, encontram-se os modelos quantitativos de suporte a decisões do processo de gestão, entre os quais consta o modelo *credit scoring*.

“O *credit scoring*, é um modelo de avaliação do crédito que recorre a uma fórmula estatística desenvolvida com base em dados de registos, financeiros, patrimoniais e de idoneidade dos clientes” (Fama & Santos, 2007). Este modelo tem sido muito utilizado no processo de gestão do risco de crédito para auxiliar na previsão de clientes

inadimplentes, visando a identificação de características que permitem distinguir os bons dos maus clientes. Neste trabalho, o modelo de *credit scoring* será então aplicado a um banco alemão na avaliação de clientes solicitantes de crédito com certas características, para determinar a que grupo pertencem: “bom crédito” (clientes adimplentes) ou “mau crédito” (clientes inadimplentes).

Clientes adimplentes são aqueles que cumprem um acordo estabelecido por um contrato, ou seja, que cumprem as suas obrigações atempadamente. Por outro lado, clientes inadimplentes são aqueles que não cumprem um contrato ou as obrigações legais dentro dos prazos previamente estabelecidos.

Pretende-se assim, com este trabalho, desenvolver um modelo *credit scoring* que, contribuindo para a identificação e resolução de problemas relacionados com o crédito, possa ser implementado no processo de análise de crédito do banco em estudo.

1.1 Motivação e Objetivos do Trabalho

A motivação deste estudo prende-se com o facto de poder aplicar um método que facilita e torna mais rápido e seguro o processo de análise de crédito. Este método pode servir como ferramenta ao especialista de crédito auxiliando-o a tomar a decisão de concessão de crédito. O objetivo deste trabalho, é o de desenvolver o modelo *credit scoring*, estimando um modelo *probit*, para análise do processo da tomada de decisão de concessão de crédito. Pretende verificar-se se a estimação do modelo a partir de uma amostra de 1000 clientes extraídos da carteira de crédito de um banco alemão contribuirá para uma melhor identificação do risco dos clientes e classificação dos mesmos de acordo com o *status* de adimplência.

1.2 Estrutura do Relatório

Este relatório está estruturado em quatro capítulos, sendo o primeiro a introdução onde se contextualiza o tema em análise, onde se apresentam a motivação e os objetivos, e a estrutura do projeto. O capítulo dois, debruça-se sobre o enquadramento teórico apresentando os conceitos de crédito, risco de crédito, especificando as causas de risco de crédito e a definição detalhada do modelo de *credit scoring*. Ainda neste capítulo são apresentadas as técnicas aplicadas em modelos de *credit scoring*, e é definido o modelo *probit*. No terceiro capítulo é apresentada a amostra e as variáveis. É ainda neste capítulo que são feitos os testes de significância para identificação das variáveis relevantes, assim como a estimação dos efeitos parciais das variáveis explicativas e a determinação do poder preditivo e a posterior interpretação dos resultados. Por último, apresentam-se as conclusões retiradas no projeto, assim como algumas sugestões para pesquisas futuras.

Capítulo 2 Enquadramento Teórico

2.1 O Conceito de Crédito

Segundo (Blatt, 1999) “crédito é crer, e crer é confiar”. “A palavra crédito tem a sua origem na palavra latina “*credere*”, que significa crer, confiar, acreditar, ou ainda, do substantivo *creditum*, o qual significa literalmente confiança”.

Desta forma, podemos dizer que o conceito de crédito está diretamente ligado a qualquer relação ou transação que se estabelece na confiança entre duas partes.

Para (Vaz, 1999) “operações de crédito bancário são aquelas em que o banco intervém como vendedor de dinheiro. Este aplica, em financiamentos e operações comerciais, os capitais captados através dos depósitos, ou seja, das operações passivas que realiza. Neste caso, o crédito é visto como uma ferramenta que gera juros e comissões”.

Considerando a exposição de (Vaz, 1999), definimos crédito como sendo o ato pelo qual o banco disponibiliza um montante em dinheiro a uma pessoa física ou jurídica, confiando que esta devolva o montante em dinheiro acrescido de juros. Estes juros são devidos pela utilização do dinheiro num determinado período acordado entre o banco e a pessoa física ou jurídica. Também se pode dizer que as operações de crédito são vistas como uma das principais funções dos bancos.

Segundo (Blatt, 1999), “existe um risco a cada concessão ou operação de crédito bancário”. O presente trabalho foca pois a vertente da concessão ou de operações de créditos bancários com o objetivo de prever a probabilidade do risco a estas inerente.

2.2. O Risco de Crédito

Pode verificar-se que em todas as operações de crédito existe um risco. Assim, antes de definirmos risco de crédito, vamos introduzir, de forma resumida, o significado da palavra risco.

Segundo (Chaia, 2003), a palavra “risco” deriva do latim *risicare*, que significa “ousar”. Este autor destaca ainda outros significados para “risco”, como “variância de resultado”, “fator catastrófico negativo”, “fator positivo de oportunidade”, bem como a relação “risco e incerteza”.

De acordo com (Blatt, 1999), “risco é a probabilidade de que algo por nós esperado e desejado não aconteça ou, complementarmente, que algo indesejado aconteça”.

Para (Neves, 2001), “o risco é a possibilidade de perdas. Assim quanto maior a possibilidade de perda, maior o risco”.

Atendendo às definições citadas podemos concluir que risco é a possibilidade de um acontecimento incerto, fortuito e de consequências negativas ou danosas.

Pretende-se de seguida definir risco de crédito.

“O risco de crédito é o risco de que as obrigações financeiras para com o credor não sejam pagas a tempo e plenamente, como esperado e contratado, resultando em um prejuízo financeiro para o credor. O risco de crédito está relacionado com o cliente, porque a dimensão do risco depende do desejo e da capacidade do cliente cumprir as suas obrigações” (Blatt, 1999).

“O risco de crédito acontece pela obrigação de resgate ou liquidação que não seja honrada pela outra parte” (Carrilho, 2009).

Já para (Bessis, 2010), “o risco de crédito é o risco da contraparte incumprir o pagamento da sua obrigação”.

De acordo com as definições apresentadas, podemos dizer que o risco de crédito é a possibilidade de pessoas físicas (indivíduos particulares) ou pessoas jurídicas (empresas) a quem o banco concedeu crédito não pagarem completamente e atempadamente o valor em dívida ou os juros correspondentes. O risco de crédito é pois a probabilidade de incumprimento do devedor para com o credor.

“No contexto bancário, a atividade tida como mais relevante e principal geradora de rendimentos é a de concessão de crédito” (Caiado & Caiado, 2009). Sendo assim, a decisão associada deve ser cuidada e muito bem analisada, antes de tomada. De facto, uma análise errada pode resultar numa má decisão e provocar grandes prejuízos financeiros. Tal pode mesmo originar uma situação de insolvência ou outros tipos de riscos inerentes à atividade bancária como, por exemplo, o risco de liquidez, que “reflete a exposição das instituições de crédito a situações de indisponibilidade de fundos perante depositantes ou credores” (Vaz, 1999).

Uma situação de insolvência na banca deve ser evitada pois pode repercutir-se fortemente a outros sectores provocando falências de empresas, instabilidade social e mesmo desempregos, o que justifica também o interesse deste trabalho.

2.3 Causas do Risco de Crédito

Para (Vaz, 1999), “o risco de crédito de uma carteira de empréstimo depende de dois grupos de fatores”:

- Fatores determinados exogenamente, tais como o estado da economia, acidentes naturais, etc.
- Fatores dependentes da capacidade de gestão interna.

Ainda que o primeiro grupo de fatores ocorra de modo exógeno, a atitude dos gestores bancários deve ser sempre positiva, tentando evitar, no máximo, práticas inseguras que originem riscos acrescidos.

O segundo grupo está relacionado com a filosofia ou atitude dos gestores perante o risco.

(Vaz, 1999) afirma que “o risco de crédito varia de acordo com os movimentos cíclicos da economia, sendo particularmente sensível nos períodos de recessão-depressão”.

(Blatt, 1999) identifica “vários fatores como podendo contribuir para aumentar o risco de crédito. Tais fatores nunca ocorrem sozinhos ou em proporções previamente definidas. Na realidade existe uma forte inter-relação e interdependência entre os diversos fatores de risco”.

2.4 Análise de Crédito

Com o conceito de crédito, verificamos que existe sempre um risco em toda operação de crédito. É de realçar que segundo (Morais, 2011), para (Casu, Girardone, & Molyneux, 2006) o risco de crédito é considerado o “principal risco da atividade bancária e o mais difícil de quantificar”. Por esta razão, as instituições financeiras preocupam-se cada vez mais com a gestão e análise de crédito.

Segundo (Blatt, 1999), “a análise de crédito consiste no estudo da situação global de um devedor em perspectiva, buscando a elaboração de um parecer que retrate, de forma clara e objetiva, o desempenho económico-financeiro do mesmo”.

Também (Santos, 2003) refere a “análise de crédito como tendo o objetivo de averiguar a compatibilidade do crédito solicitado com a capacidade financeira do cliente”.

Assim, o processo de análise de crédito, procura apurar os dados mais relevantes que influenciam a tomada de decisão para concessão ou não de crédito. Para tal, utilizam-se as técnicas ou procedimentos mais adequados para a análise de crédito.

(Fama & Santos, 2007) “indicam que o processo de análise e concessão de crédito recorre ao uso de duas técnicas: a técnica subjetiva e a técnica objetiva ou estatística. A primeira diz respeito à técnica baseada no julgamento humano e a segunda é baseada em processos estatísticos”.

Seguindo ainda os mesmos autores, podemos perceber que, a primeira técnica está relacionada com a habilidade e experiência profissional adquirida pelo analista ou agente de crédito. Esta é uma técnica em que o especialista em crédito se baseia no seu conhecimento técnico, nas informações disponíveis, na habilidade em tomar uma decisão de crédito e na sensibilidade humana, ou seja, é uma técnica baseada apenas no julgamento do especialista em crédito. A segunda técnica está relacionada com os processos estatísticos ou matemáticos, ou seja, envolve uma análise mais quantitativa para a tomada da decisão de crédito.

De forma resumida vamos fazer a descrição de cada uma das técnicas de análise de crédito acima mencionadas. Pretende-se, assim, melhor compreender a importância de cada uma face à exposição ao risco.

2.4.1 Análise Subjetiva do Crédito

“A análise subjetiva, é baseada na experiência adquirida dos analistas de crédito, no conhecimento técnico, no bom senso e na disponibilidade de informações (internas e externas) que lhes possibilitem diagnosticar se o cliente possui idoneidade e capacidade de gerar receita para honrar o pagamento das parcelas dos financiamentos” (Santos, 2003).

Nesta análise, é feito o estudo das informações disponíveis, relacionadas com a capacidade financeira e idoneidade do cliente. Essas informações permitem ao analista de crédito identificar fatores de risco que possam comprometer a capacidade de pagamento do cliente.

Segundo (Silva, 2011), “na análise subjetiva de crédito, o analista da área de crédito agrupa as informações do cliente baseando-se nos 5 C’s do crédito (mesmo que implicitamente) ou seja: Caráter, Capacidade, Capital, Condições e Colateral. Porém, além destes C’s tradicionais há também um outro C, utilizado em algumas situações específicas, o Conglomerado”.

- **Caráter:** analisa a honestidade, a moral e intenção do cliente em honrar o seu compromisso;
- **Capacidade:** analisa as qualidades e habilidades profissionais do cliente, em converter os seus negócios em renda/receita. (Mantovani, Gonçalves, & Gouvêa, 2013) consideram a “capacidade o aspeto mais subjetivo, pois depende mais da perceção do analista do que da análise de dados registados”;

- **Capital:** analisa a situação financeira e a capacidade de pagamento do cliente, a sua fonte de receita e os recursos financeiros, para comprovar se o cliente tem meios de liquidar o empréstimo;
- **Condição:** analisa o impacto de fatores externos que podem influenciar a operação, como a conjuntura económica do país que, de certa forma, pode afetar o compromisso do cliente diante das suas obrigações;
- **Colateral:** para este C são analisadas as garantias disponibilizadas pelo cliente para compensar ou reforçar a operação de crédito. Contudo, (Silva, 2011) “ressalta que uma garantia não pode justificar uma decisão de concessão de crédito para um cliente cujo risco seja extremamente elevado”;
- **Conglomerado:** “análise não apenas de uma empresa específica que esteja pleiteando o crédito, mas do grupo económico do qual a empresa faz parte, isto é, de todas as empresas de um mesmo proprietário ou empresa controladora” (Silva, 2011).

“A análise dos C’s de crédito deve ser feita de forma conjunta, ou seja, o especialista em crédito deve incorporar todos os C’s de crédito no processo de tomada de decisão. Por exemplo, de nada adianta um tomador de crédito ter vontade de pagar (ser honesto) se ele não tem capacidade, ou seja, se o empréstimo for maior que o seu rendimento” (Silva, 2011).

Embora esta técnica mostre ser muito utilizada pelos analistas de crédito, ela possui muitas desvantagens, tais como: torna-se exaustiva e difícil para o especialista em crédito examinar todos os registos dos solicitantes de crédito, sendo ele responsabilizado pelas falhas, caso haja; é muito propensa a falhas humanas; podem existir várias opiniões num

só relatório, isto é, incoerência de análises atendendo ao facto de haver decisões divergentes entre diferentes analistas para a concessão de crédito; as fraudes também constituem um risco neste tipo de técnica ou análise.

A análise subjetiva, como o nome indica, é uma técnica que se baseia muito no critério humano, não é uma ciência exata, o que faz com que haja falhas. Tendo em conta este pormenor, com intenção de melhorar este quadro, foi desenvolvida a análise objetiva (quantitativa).

2.4.2 Análise Objetiva de Crédito

Esta análise tem como foco os métodos estatísticos que permitem encontrar resultados que certifiquem a capacidade de pagamento do solicitante de crédito. Nesta análise são criadas pontuações/*scores* de crédito que servem de suporte para o analista, na previsão da probabilidade de um cliente se tornar incumpridor. De entre as técnicas objetivas existentes para gestão do risco de crédito, destacamos o modelo de *credit scoring*.

2.4.3 Credit Scoring

Trata-se de um modelo de avaliação do crédito que estima a probabilidade de haver incumprimento. Para (Fama & Santos, 2007), “os precursores da metodologia do *credit scoring* para identificação de bons e maus tomadores de financiamento foram o Fisher (1936) e Durand (1941)”.

Para o estudo estatístico, os bancos recolhem e selecionam as informações mais relevantes em termos de registos dos clientes. De seguida, são atribuídos pesos ou ponderações de acordo com os critérios de importância associados a políticas internas de crédito do banco. Gera-se assim um sistema de pontuação/*score* para cada cliente.

Segundo (Fama & Santos, 2007), define-se uma pontuação mínima para a soma das pontuações/*score* de todos os clientes de uma carteira que indica o ponto de corte ou ponto de distinção de clientes. Este ponto é o requisito mínimo para aprovação ou recusa do crédito. Portanto, um cliente com um *score* maior que o ponto de corte terá o crédito aprovado. Contrariamente, caso o *score* seja inferior ao ponto de corte, o crédito é recusado.

O foco principal do modelo *credit scoring* é identificar os solicitantes de créditos de acordo com a sua probabilidade de inadimplência, ou seja, o seu objetivo é o de prever a probabilidade de um cliente, com determinadas características, ser considerado como um potencial risco. Assim, o *credit scoring* visa identificar características ou fatores fundamentais que influenciam a adimplência ou inadimplência dos clientes, possibilitando a classificação dos mesmos em grupos distintos e, conseqüentemente, a tomada de decisão sobre a aceitação ou não do crédito em análise.

O *credit scoring*, mesmo assente num processo científico, “não inibe a possibilidade de se recusar um bom pagador ou se aceitar um mau pagador”, (Sousa & Chaia, 2000) citado por (Neto & Carmona, 2004).

Ainda em termos de classificação de tipos de clientes, (Mantovani, Gonçalves, & Gouvêa, 2013) referem que “há um custo mais elevado em se classificar erroneamente um mau pagador do que um bom pagador, isto é, há maior dano para a organização se um mau cliente for classificado como bom pagador do que a situação contrária”.

O modelo *credit scoring* é pois uma boa base de avaliação do risco de crédito, pois automatiza o processo de crédito tornando-o mais ágil e reduzindo as percentagens de

inadimplência. Contudo, como qualquer outro modelo, este também apresenta vantagens e desvantagens.

(Neto & Carmona, 2004), apresentam as principais vantagens e desvantagens do modelo *credit scoring* de acordo com (Parkinson & Ochs, 1998), que se resumem de seguida:

Vantagens do modelo *credit scoring*:

- **Informações organizadas:** A sistematização e organização das informações contribuem para a melhoria do processo;
- **Compreensão do processo:** O modelo desenvolvido sintetiza o processo de concessão de crédito da instituição, fornecendo maiores subsídios para entendê-lo;
- **Maior eficiência do processo:** A análise de crédito é centrada em um número menor de fatores, reduzindo o tempo do processo e melhorando a eficiência.

Desvantagens do modelo *credit scoring*:

- **Custo de desenvolvimento:** o desenvolvimento de um sistema de *credit scoring* traz não apenas os custos do sistema, mas altos custos de manutenção com suporte técnico e recursos humanos;
- **Modelos com “excesso de confiança”:** Os modelos de *credit scoring* podem adquirir Status de “perfeitos”, não permitindo crítica de seus resultados;
- **Interpretação equivocada dos *escores*:** O uso inadequado do sistema devido à falta de conhecimento pode ocasionar problemas para a instituição.

2.5 Técnicas Estatísticas aplicadas em Modelos de *Credit Scoring*

Segundo (Neto & Carmona, 2004), “as três principais técnicas para elaboração de sistemas de *credit scoring* são a análise discriminante, a regressão logística e, mais recentemente modelos baseados em redes neuronais. A análise discriminante foi a primeira técnica estatística utilizada na criação de modelos de *credit scoring*. A aplicação de modelos matemáticos e estatísticos para a decisão de concessão de crédito ganhou reconhecimento quando David Duran, da *National Bureau of Economic Research* (Agência Nacional de Pesquisa Económica), em Nova Iorque, publicou o seu estudo em 1941 intitulado “*Risk Elements in Consumer Installment Financing*” (Elementos de Risco no Financiamento a Prazo do Consumidor). Duran foi o primeiro autor a usar a análise discriminante para medir o risco de crédito”.

De acordo com (Silva, 2011), “uma grande variedade de técnicas estatísticas e de investigação operacional foi já utilizada no desenvolvimento de modelos de *credit scoring* tais como análise discriminante, regressão logística, árvores de decisão, redes neuronais artificiais, algoritmos genéticos e programação matemática”.

De acordo com os autores, percebemos que existem outras técnicas, para além das técnicas estatísticas, que são utilizadas no desenvolvimento de modelos de *credit scoring*. Neste trabalho optou-se por avaliar a probabilidade de um cliente do banco em estudo ser adimplente ou inadimplente utilizando a técnica estatística conhecida como regressão *probit*.

2.6 O Modelo *Probit*

O modelo *probit* é um modelo de regressão para variável de resposta binária. A variável dependente é qualitativa e é classificada em duas categorias que assumem respetivamente valor 1 no caso de acontecer um evento e 0 caso contrário. Esta regressão é utilizada para analisar a relação entre um conjunto de particularidades de uma observação e a probabilidade de que esta observação pertença a um dos dois grupos identificados. É comum o uso desta regressão para análise de dados com resposta binária ou dicotómica. No presente trabalho temos:

$$Y = \begin{cases} 1, & \text{se um cliente é classificado adimplente} \\ 0, & \text{se um cliente é classificado inadimplente} \end{cases}$$

O modelo *probit* é expresso como:

$$P(Y = 1|\mathbf{x}) = F(\mathbf{x}'\boldsymbol{\beta}) = \Phi(\mathbf{x}'\boldsymbol{\beta}) \quad [1]$$

Onde $\Phi(\mathbf{x}'\boldsymbol{\beta})$ é função de distribuição normal estandardizada, na qual $\boldsymbol{\beta}$ representa o vetor de parâmetros associados às variáveis e \mathbf{x} o vetor de variáveis independentes. O *probit* é um modelo estimado pelo método da máxima verosimilhança.

No âmbito da análise de crédito, o *probit* é utilizado para a avaliação da probabilidade de inadimplência de clientes em situações relacionadas à concessão de crédito.

O modelo *probit* permite analisar a relação entre as variáveis independentes e a probabilidade de adimplência.

Para se encontrar o efeito parcial de variáveis contínuas ou discretas sobre a probabilidade de resposta, devemos ter em consideração:

- Se X_j é uma variável contínua, o seu efeito parcial sobre a probabilidade de resposta é obtido a partir da derivada parcial:

$$\frac{\partial \Phi(\cdot)}{\partial X_j} = \beta_j \cdot \varphi(\mathbf{x}'\boldsymbol{\beta}) \quad [2]$$

onde φ é a função densidade de probabilidade da normal estandardizada.

- Se X_k é uma variável *dummy*, então o seu efeito parcial sobre a probabilidade de resposta é dado por:

$$\Phi(\beta_0 + \beta_1 X_1 + \dots + \beta_{k-1} X_{k-1} + \beta_k) - \Phi(\beta_0 + \beta_1 X_1 + \dots + \beta_{k-1} X_{k-1}) \quad [3]$$

O cálculo do efeito parcial médio de X_j é dado por:

$$EPM_j = \hat{\beta}_j \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{x}'_i \hat{\boldsymbol{\beta}}) \quad [4]$$

Enquanto o efeito parcial médio de X_k é igual a:

$$EPM_k = \frac{1}{n} \sum_{i=1}^n [\Phi((\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_{k-1} X_{ik-1} + \hat{\beta}_k) - \Phi(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \dots + \hat{\beta}_{k-1} X_{ik-1}))] \quad [5]$$

Capítulo 3 Aplicação Empírica

3.1 Apresentação da Amostra e das Variáveis

A amostra utilizada neste trabalho é composta por 1000 clientes, extraídos da carteira de crédito de um banco alemão¹. Para cada cliente a variável de resposta binária “credibilidade” é observável, sendo também consideradas 20 variáveis independentes que caracterizam o cliente e o tipo de crédito, a maioria de tipo qualitativo (categóricas). Na base de dados da carteira de créditos do banco, observam-se dois grupos distintos de tipos de crédito, classificados em “bom crédito” e “mau crédito”. A tabela do Anexo 1 contém

¹ Dados disponíveis em <http://www.ipipan.eu/~teisseyp/TEACHING/ZMUM/DANE/kredit.htm>

o resumo das variáveis da base de dados. Dos 1000 clientes da carteira de crédito, 700 pertencem ao grupo “bom crédito” (clientes adimplentes) e os restantes 300 pertencem ao grupo “mau crédito” (clientes inadimplentes). Esta classificação foi feita pelos especialistas bancários de acordo com os padrões da instituição.

Sendo a maioria das variáveis independentes do tipo qualitativo (categóricas), torna-se necessário a sua transformação em *dummies*, para ser possível a sua representação numérica, assumindo valor 1 para indicar a presença de um atributo e 0 para indicar a ausência desse atributo. Consideram-se $k-1$ *dummies* para cada variável qualitativa de k categorias. Porém, devido ao facto de existirem variáveis qualitativas com muitas categorias, foi necessário o agrupamento das categorias que tratam de forma indireta ou direta do mesmo assunto. Por exemplo, a variável “Objetivo do Crédito” possui 11 categorias: *Carro novo*; *Carro usado*; *Artigos mobiliários*; *Rádio / televisão*; *Eletrodomésticos*; *Reparações*; *Educação*; *Férias*; *Reeducação*; *Negócios* e *Outros*. Estas 11 categorias desta variável foram agrupadas em cinco: *Carro novo* e *Carro usado* deram origem à variável *dummy* “auto”; das categorias *Educação* e *Reeducação* formulámos a variável *dummy* “educ” e assim procedemos com todas as variáveis que apresentam as mesmas características, como se ilustra na Tabela 1.

A variável dependente, designada por **CRED**, é uma variável binária que assume o valor 1 se o cliente for classificado como adimplente. Para explicar a probabilidade de adimplência de um cliente, consideram-se 13 variáveis independentes. O nome das variáveis assim como a sua definição estão descritas na Tabela 1.

Variável	Descrição	Definição
AD	Ativos Disponíveis - variável categórica com a seguinte classificação: 1. Sem ativos disponíveis; 2. Carro/outro; 3. Contrato de poupança com uma sociedade de construção/seguro de vida; 4. Posse de casa ou terra. Definida uma <i>dummy</i> ,	Conjunto de bens e direitos do solicitante do crédito.
PCT	= 1 se o solicitante do crédito possui casa ou terra e 0 caso contrário.	
CRED	Credibilidade - variável dependente binária = 1 se o cliente é adimplente e 0 caso contrário.	Classificação do crédito em bom (merecedor de crédito) ou mau (não merecedor de crédito), pelos especialistas bancários.
DC	Duração do Crédito (em meses).	Prazo estabelecido pelo banco em que perdura o pagamento integral do crédito.
IDADE	Idade (em anos).	Número de anos do solicitante de crédito.
MC	Montante do Crédito (em <i>Deutsche Mark</i>).	É a quantia em valor que o banco cede ao solicitante do crédito, caso for aprovado.
NCANB	Número de Créditos Anteriores Neste Banco (incluindo o que está em execução) - variável categórica com a seguinte classificação: 1. 1 Crédito; 2. 2 ou 3 créditos; 3. 4 ou 5 créditos; 4. 6 ou mais créditos. Definida uma <i>dummy</i> ,	Quantidade de empréstimos tomados anteriormente num banco, incluindo o atual.
NCm6	= 1 se o número de créditos for maior ou igual a 6 e 0 caso contrário.	
ND	Número de Dependentes - variável categórica com a seguinte classificação: 1. 3 ou mais pessoas; 2. 0 a 2 pessoas. Definida uma <i>dummy</i> ,	Número de pessoas que são sustentadas pelo solicitante do crédito.
NDm3	= 1 se o número de dependentes for maior ou igual a 3 e 0 caso contrário.	
OC	Objetivo do Crédito - variável categórica com a seguinte classificação: 0. Outros; 1. Carro novo; 2. Carro usado; 3. Artigos mobiliários; 4. Rádio/televisão; 5. Eletrodomésticos; 6. Reparações; 7. Educação; 8. Férias; 9. Reeducação; 10. Negócios. Definidas 5 <i>dummies</i> , com grupo base Outros .	Intenção ou finalidade de obter o crédito para satisfação de uma necessidade.
Auto	= 1 se o objetivo do crédito for a compra de um carro novo ou usado e 0 caso contrário.	
Casa	= 1 se o objetivo do crédito for a compra de eletrodomésticos ou reparações e 0 caso contrário.	
Educ	= 1 se o objetivo do crédito for para educação ou reeducação e 0 caso contrário.	
Fer	= 1 se o objetivo do crédito for para férias e 0 caso contrário.	
Nego	= 1 se o objetivo do crédito for para negócios e 0 caso contrário.	

Variável	Descrição	Definição
OCE	Outros Créditos em Execução - variável categórica com a seguinte classificação: 1. Em outros bancos; 2. Na loja de departamentos ou na casa da ordem de correio; 3. Sem mais créditos em execução. Definidas duas <i>dummies</i> .	Existência de outras obrigações financeiras para cumprir.
OB	= 1 se existe outros créditos em outros bancos e 0 caso contrário.	
LD	= 1 se existe outros créditos na loja de departamentos e 0 caso contrário	
OCUP	Ocupação - variável categórica com a seguinte classificação: 1. Desempregado/não qualificado sem residência permanente; 2. Empregado não qualificado com residência permanente; 3. Funcionário público excluindo quadros superiores; 4. Executivo/por conta própria/quadro superior da função pública. Definidas duas <i>dummies</i> .	Atividade diária que gera receita ao solicitante do crédito.
FPEQS	= 1 se o solicitante do crédito é funcionário público não quadro superior e 0 caso contrário.	
E/CP/QSFP	= 1 se o solicitante do crédito é funcionário executivo /por conta própria /quadro superior da função pública e 0 caso contrário.	
SCC	Saldo de Conta Corrente - variável categórica com a seguinte classificação: 1. Sem movimento na conta; 2. Sem saldo ou débito; 3. $0 \leq \dots < 200$ DM; 4. ≥ 200 DM. Definidas 3 <i>dummies</i> .	Movimentos validados até ao momento da consulta e do qual se pode dispor de imediato.
SMC	= 1 se o solicitante não tem movimento na conta e 0 caso contrário.	
SSD	= 1 se o solicitante não tem saldo na conta e 0 caso contrário.	
$0 \leq \dots < 200$	= 1 se o saldo do solicitante está entre 0 e 200 e 0 caso contrário.	
SPCE	Situação perante Créditos Existentes - variável categórica com a seguinte classificação: 1. Falhas no pagamento de créditos anteriores; 2. Conta corrente problemática/há mais créditos em execução em outros bancos; 3. Sem créditos anteriores/pagou todos créditos anteriores; 4. Sem problemas com créditos atuais neste banco; 5. Créditos anteriores pagos neste banco. Definidas duas <i>dummies</i> .	Liquidação dos empréstimos tomados anteriormente.
FPCA	= 1 se houve falhas no pagamento de créditos anteriores e 0 caso contrário	
MCEOB	= 1 se existem mais créditos em execução em outros bancos e 0 caso contrário.	

Variável	Descrição	Definição
TAXA	Prestação em % do rendimento disponível - variável categórica com a seguinte classificação: 1. ≥ 35 ; 2. $25 \leq \dots < 35$; 3. $20 \leq \dots < 25$; 4. < 20 . Definidas 3 <i>dummies</i> .	Prestações de pagamentos do crédito durante o prazo estabelecido.
Prestm35	= 1 se a prestação do pagamento do crédito for maior ou igual a 35% do rendimento e 0 caso contrário	
Prestm25 e 35	= 1 se a prestação do pagamento do crédito for maior ou igual a 25% e inferior a 35% do rendimento e 0 caso contrário.	
Prestm20 e 25	= 1 se a prestação do pagamento do crédito for maior ou igual a 20% e inferior a 25% do rendimento e 0 caso contrário.	
TE	Tempo de Emprego - variável categórica com a seguinte classificação: 1. Desempregado; 2. ≤ 1 ano; 3. $1 \leq \dots < 4$ anos; 4. $4 \leq \dots < 7$; 5. ≥ 7 . Definidas duas <i>dummies</i> , com grupo base o ponto 3 ($1 \leq \dots < 4$)	Anos que um indivíduo está a trabalhar em determinado emprego.
Desemp	= 1 se o solicitante do crédito está desempregado e 0 caso contrário.	
Empm4	= 1 se o solicitante do crédito está empregado há 4 ou mais anos e 0 caso contrário.	

Tabela 1 Definição das Variáveis

Para o estudo das variáveis selecionadas vamos recorrer à estatística descritiva e à análise de correlação.

Estatística descritiva

De acordo com os resultados da tabela do Anexo 2 observou-se que existe uma grande dispersão das variáveis. Verificou-se também que, em média, dos créditos solicitados: 28,4% são para automóveis; 36,5% são para casa (compra de eletrodomésticos ou reparações); 70% são considerados adimplentes. A duração média do crédito é aproximadamente igual a 21 meses, sendo o mínimo de 4 e o máximo de 72 meses. Em média, 42,7% dos indivíduos que solicitam crédito, são empregados há 4 ou mais anos. A idade média está próxima de 36 anos, variando entre 19 (mínimo) e 75 (máximo). O

montante médio do crédito é de 3.271,248 *deutsche mark*, sendo o mínimo de 250 e o máximo de 18.424 *deutsche mark*.

Análise de correlação

Analisando os resultados da matriz de correlação (Anexo 3), podem retirar-se as conclusões que se apresentam de seguida.

A variável dependente **CRED** está positivamente correlacionada com as variáveis **IDADE, auto, casa, fer, prestm35, prestm25_a_35, prestm20_a_25, empm4, FPEQS** e **de_0_a_200**. Atendendo a correlação positiva destas variáveis, significa que estas exercem uma influência positiva na probabilidade do crédito ser aprovado.

Contrariamente as correlações positivas, verificamos também que **CRED** está negativamente correlacionada com as variáveis **MC, DC, Educ, Nego, OB, LD, FPCA, MCEOB, NCm6, Desemp, QSFP, PCT, NDm3, SMC** e **SSD**. A correlação negativa destas variáveis, significa que exercem uma influência negativa sobre a probabilidade do crédito ser aprovado.

Das 20 variáveis existentes na base de dados, seleccionámos 13. Consideramos como critérios de seleção das variáveis, a sua importância económica no processo da concessão de crédito e a correlação entre a variável dependente e as independentes. Tendo em conta a seleção das variáveis, torna-se necessário fazer o estudo do sinal esperado das mesmas, como se explica de seguida.

Montante do crédito (MC): espera-se um sinal negativo para esta variável, isto é, quanto maior for o montante, maior poderá ser a dificuldade em pagá-lo;

Auto: espera-se um sinal positivo para esta variável, que integra um tipo de crédito (crédito automóvel) muito solicitado nas carteiras de créditos bancários. Esta variável refere-se a um bem imobilizado que o banco pode tomar de volta em caso de incumprimento;

Para as variáveis **Casa**, **Educ**, **Fer** e **Nego** espera-se um sinal negativo para cada uma, visto que as mesmas se enquadram no tipo de crédito pessoal, que tem um taxa de juro mais alta (Deutsche Bank, 2017);

Duração do crédito (DC): espera-se um sinal negativo, ou seja, quanto mais tempo perdurar o pagamento do empréstimo solicitado, maior é a probabilidade de o cliente se tornar inadimplente;

Prestm35, **Prestm25_e_35** e **Prestm20_e_25:** espera-se um sinal negativo para cada variável, tendo em conta que representa um esforço de pagamento superior ao do grupo base, sendo assim, haverá maior dificuldade em pagar o crédito;

Outros bancos (OB) e Loja de departamentos (LD): espera-se um sinal negativo para estas variáveis, visto que representam outros encargos financeiros a cumprir;

NCm6: espera-se um impacto negativo para esta variável, visto que um número de créditos solicitados maior ou igual a seis leva a muitas obrigações financeiras para cumprir, dificultando assim o pagamento das mesmas;

FPCA: espera-se um impacto negativo para esta variável, tendo em conta que as falhas de pagamentos de créditos anteriores demonstram o não cumprimento de obrigações contratuais no processo de concessão de crédito;

MCEOB: esta variável também indica a existência de outros encargos financeiros para cumprir o que faz com que o seu sinal esperado seja negativo;

IDADE: é esperado um sinal positivo para esta variável tendo em conta a maturidade e responsabilidade que se adquire com certa idade.

Desemp: espera-se um sinal negativo para esta variável tendo em conta que os indivíduos desempregados não apresentam garantias de fonte de rendimentos;

Empm4: espera-se um sinal positivo para esta variável visto que a duração do emprego pode determinar a capacidade financeira do cliente. Ou seja, quanto maior for o tempo de emprego de um indivíduo, maior confiança inspira ao banco e maior capacidade financeira pode adquirir;

FPEQS e QSFP: espera-se um sinal positivo para estas variáveis, porque identificam os indivíduos assalariados. Ou seja, todos os solicitantes de créditos que têm uma ocupação e esta constitui a sua fonte de receita, têm maior probabilidade de serem classificados como adimplentes;

PCT: espera-se um sinal positivo para esta variável, visto que as garantias funcionam como a vinculação de um bem que assegure a liquidação do crédito caso o tomador não honre os seus compromissos;

NDm3: espera-se um sinal negativo para esta variável, tendo em conta que os indivíduos com maior número de dependentes, também têm maiores obrigações/responsabilidades financeiras;

Smc, Ssd e De 0_a_200: espera-se um sinal negativo para estas variáveis, visto que podem caracterizar o comportamento financeiro do solicitante de crédito. Ou seja, quanto menos forem os movimentos e menores forem os saldos bancários de um cliente, maior é a probabilidade de este ser classificado como inadimplente.

3.2 Estimação

Para estimação do modelo *probit* utilizou-se o *software Eviews*, com a referida amostra de 1000 clientes divididos em duas categorias, 700 clientes adimplentes e 300 clientes inadimplentes. Às 13 variáveis independentes selecionadas, juntaram-se as variáveis *dummy* criadas a partir das variáveis categóricas, resultando num total de 25 variáveis

incluídas no modelo. A relação entre a variável dependente e as variáveis independentes é dada pela fórmula [1].

A Tabela 2 apresenta a estimação dos coeficientes dos modelos considerados.

Variável	Modelo 1			Modelo 2		
	Coefficiente	Erro padrão	Significância	Coefficiente	Erro padrão	Significância
Constante	0,973	0,6715	0,147	0,941	0,209	0
LOG(MC)	-0,019	0,094	0,841			
IDADE	0,011	0,005	0,016	0,010	0,004	0,019
DC	-0,020	0,005	0,000	-0,020	0,004	0
AUTO	0,458	0,129	0,000	0,344	0,115	0,003
CASA	0,328	0,123	0,008	0,224	0,110	0,041
EDUC	0,300	0,178	0,092			
FER	0,918	0,617	0,136			
NEGO	0,635	0,410	0,121			
PRESTM35	0,350	0,162	0,031	0,278	0,142	0,051
PRESTM25_A_35	0,228	0,126	0,071	0,194	0,113	0,086
PRESTM20_A_25	0,141	0,138	0,309			
OB	-0,197	0,136	0,149			
LD	-0,204	0,216	0,344			
FPCA	-0,644	0,227	0,005	-0,598	0,222	0,007
MCEOB	-0,443	0,211	0,036	-0,514	0,196	0,009
NCM6	-0,147	0,604	0,808			
DESEMP	0,001	0,197	0,996			
EMPM4	0,318	0,104	0,002	0,315	0,098	0,001
FPEQS	-0,028	0,120	0,814			
QSFP	-0,092	0,171	0,591			
PCT	-0,281	0,136	0,039	-0,317	0,130	0,015
NDM3	0,009	0,132	0,945			
SMC	-1,122	0,119	0	-1,125	0,117	0
SSD	-0,800	0,121	0	-0,783	0,119	0
DE_0_A_200	-0,471	0,204	0,021	-0,468	0,202	0,020
McFadden R-squared	0,190			McFadden R-squared	0,182	
Akaike info criterion	1,041			Akaike info criterion	1,028	
LR statistic	232,2451			LR statistic	221,811	
Prob(LR statistic)	0			Prob(LR statistic)	0	
Log likelihood	-494,742			Log likelihood	-499,959	

Tabela 2 Modelos estimados

De notar que, apesar da seleção **inicial** das variáveis se basear na correlação entre a variável dependente e as independentes e na importância económica das variáveis sobre

o processo da concessão de crédito, verificamos que nem todas as variáveis são estatisticamente significativas, como podemos observar na Tabela 2 - Modelo 1. Sendo assim, vamos testar a significância conjunta destas variáveis assim como a significância global do modelo.

3.3 Testes de Significância Conjunta no Modelo *Probit*

Para a realização de testes de significância conjunta no modelo *probit*, recorreremos ao teste do rácio de verosimilhança – *Likelihood Ratio tests* (LR), com uma distribuição assintótica do Qui-Quadrado, com graus de liberdade (q) igual ao número das restrições que estão a ser testadas. A estatística LR é representada pela seguinte fórmula:

$$LR = 2 (L_{nr} - L_r) \sim \chi^2_{(q)} \quad [7]$$

onde q é o número de restrições sob a hipótese nula, L_{nr} é o valor do logaritmo da verosimilhança para o modelo não restrito e L_r é o valor do logaritmo da verosimilhança para o modelo restrito, $L_{nr} \geq L_r$. LR é não negativo e geralmente estritamente positivo. Ao calcular a estatística LR para modelos de resposta binária, é importante saber que a função logaritmo da verosimilhança é sempre um número negativo [para uma análise mais ampla aconselha-se Wooldridge (2009, Capítulo 17)].

De acordo com o teste LR passamos então para o teste de significância global, considerando como nível de significância 5%.

$$LR = 232,2451$$

$$P_{obs} = 0.000000 < 0,05$$

Os resultados do teste levam-nos a concluir que há evidência de significância global do Modelo 1.

No entanto, apesar de existir significância global no Modelo 1, é necessário testar a significância conjunta das variáveis que não apresentam relevância estatística individual. Iremos considerar as variáveis não relevantes a um nível de significância de 10%.

São testadas conjuntamente as variáveis *dummy* que pertencem à mesma variável qualitativa. As hipóteses e os resultados dos testes encontram-se no Anexo 4. A tabela 3 apresenta os valores dos testes de significância conjunta, assim como os valores do teste de todas variáveis não relevantes estatisticamente.

Teste de significância conjunta			
Variáveis	LR	P-Value	Significância
Fer e Nego	4,946	0,084	0,05
Ob e Ld	2,588	0,274	0,05
Fpeq e qsf	0,292	0,864	0,05
Teste global de todas variáveis não significativas			
Variáveis	LR	P-Value	Significância
Log(mc); Fer; Nego; Prestm20_a_25; Ob; Ld; Ncm6; Desemp; Fpeq; Qsf; Ndm3.	8,934	0,628	0,05

Tabela 3 Teste de significância

Feitos os testes de significância conjunta, especificou-se um novo modelo. No entanto, verificando que a variável **Educ** não era estatisticamente significativa, esta foi retirada, conduzindo ao Modelo 2 (Tabela 2).

De acordo com a estimação do Modelo 2, concluímos que há evidência de significância global e de significância individual a 5% de todas as variáveis, com exceção das variáveis **Prestm35** e **Prestm25_a_35** que são significativas a 10%.

3.4 Interpretação dos Resultados

Sendo o *probit* um modelo de resposta binária, apenas podemos interpretar o sinal dos coeficientes, não o seu valor. Para ser possível a interpretação dos coeficientes é necessário recorrer ao cálculo dos efeitos parciais. Porém, vamos antes proceder à interpretação dos sinais dos coeficientes do Modelo 2, como apresenta a Tabela 4.

Variável	Impacto	Estatisticamente significativa?	Observações
IDADE	Positivo	Sim	De acordo com o sinal esperado, a Idade exerce um impacto positivo na probabilidade de aprovação do crédito.
DC	Negativo	Sim	Em anuência com o sinal esperado, o Dc exerce um impacto negativo na probabilidade de aprovação do crédito.
AUTO	Positivo	Sim	Em anuência com o sinal esperado.
CASA	Positivo	Sim	Contrariamente ao sinal esperado, a Casa exerce um impacto positivo na probabilidade de aprovação do crédito.
PRESTM35	Positivo	Sim	Contrariamente ao sinal esperado, a Prestm35 exerce um impacto positivo na probabilidade de aprovação do crédito.
PRESTM25 A 35	Positivo		Contrariamente ao sinal esperado, a Prestm25 a 35 exerce um impacto positivo na probabilidade de aprovação do crédito.
FPCA	Negativo	Sim	Em anuência com o sinal esperado.
MCEOB	Negativo	Sim	Em anuência com o sinal esperado.
EMPM4	Positivo	Sim	Em anuência com o sinal esperado.
PCT	Negativo	Sim	Contrariamente ao sinal esperado, a Pct exerce um impacto negativo na probabilidade de aprovação do crédito. Tal na realidade, não faz sentido, visto que em muitos contratos de créditos são exigidas garantias.
SMC	Negativo	Sim	Em anuência com o sinal esperado.
SSD	Negativo	Sim	Em anuência com o sinal esperado.
DE_0__A__200	Negativo	Sim	Em anuência com o sinal esperado.

Tabela 4 Interpretação dos sinais

De modo geral, verificamos que as variáveis independentes selecionadas podem contribuir para a qualidade das previsões em termos da classificação dos clientes. Relativamente à qualidade do ajustamento do Modelo 2, de acordo com o *McFadden R – squared*, verificamos que as variáveis que o compõem conseguem apenas explicar 18% da variação total amostral da “credibilidade”. Apesar da

significância global do Modelo 2, registou-se um nível baixo de ajustamento. Tal pode ocorrer devido à omissão de variáveis. No entanto, este valor é usual para este tipo de dados.

3.5 Estimação dos Efeitos Parciais

Como vimos na secção 3.4, os coeficientes não podem ser interpretados diretamente. Os efeitos parciais das variáveis independentes não são contantes, o sinal (+) ou (-) dá o sentido do efeito. Para que interpretação seja possível vamos recorrer ao cálculo dos efeitos parciais das variáveis incluídas no Modelo 2 (Tabela 2).

De seguida analisam-se os resultados do cálculo dos efeitos parciais. Para as variáveis contínuas utilizou-se a fórmula [2] e para as variáveis discretas utilizou-se a fórmula [3]. Para obtenção dos efeitos parciais, utilizamos a função *command* presente no *software Eviews*. Os resultados assim como a sua interpretação encontram-se na Tabela 5, tendo em conta a condição *ceteris paribus* (mantendo tudo resto constante).

Variáveis contínuas	EPM	Observação: <i>Ceteris paribus</i>
Idade	0,003	Aumentando a idade em um ano, provoca um aumento médio de 0,003 na probabilidade de ser adimplente.
Dc	-0,006	Aumentando a duração do crédito em um mês, a probabilidade de ser adimplente diminui, em média, 0,006.
Variáveis discretas (dummy)	EPM	
Auto	0,093	A probabilidade de um cliente que solicita um crédito automóvel, ser adimplente aumenta, em média, 0,093 relativamente a outros créditos.
Casa *	0,062	A probabilidade de um cliente que solicita um crédito pessoal, ser adimplente aumenta, em média, 0,062 relativamente a outros créditos.
Prestm35 *	0,074	Se o cliente está num escalão de maior ou igual a 35% da prestação do pagamento, a probabilidade de ser adimplente aumenta, em média, 0,074 relativamente ao escalão de menor de 20%.
Prestm25_a_35 *	0,053	Se o cliente está num escalão de 25 a 35% da prestação do pagamento, a probabilidade de ser adimplente aumenta, em média, 0,053 relativamente ao escalão de menor de 20%.
Fpca	-0,186	Se existiram falhas de pagamentos de créditos anteriores, a probabilidade de ser adimplente diminui, em média, 0,186 relativamente a sem créditos /pagou todos créditos.

Variáveis contínuas	EPM	Observação: <i>Ceteris paribus</i>
Mceob	-0,158	Se existem mais créditos em execução em outros em bancos, a probabilidade de ser adimplente diminui, em média, 0,158 relativamente a sem créditos /pagou todos créditos.
Empm4	0,088	Se o solicitante do crédito está empregado há 4 ou mais anos, a probabilidade de ser adimplente aumenta, em média, 0,088 relativamente ao que está empregado a menos de 4 anos.
Pct *	-0,094	Se o solicitante do crédito possui ativos disponíveis (casa ou terra), a probabilidade de ser adimplente diminui, em média, 0,094 relativamente ao solicitante que não possui ativos.
Smc	-0,35	Se o solicitante do crédito está sem movimentos na conta, a probabilidade de ser adimplente diminui, em média, 0,35 relativamente ao solicitante com saldo maior que 200.
Ssd	-0,231	Se o solicitante do crédito está sem saldo, a probabilidade de ser adimplente diminui, em média, 0,231 relativamente ao solicitante com saldo maior que 200.
De_0_a_200	-0,14	Se o saldo do solicitante do crédito varia de 0 a 200 u.m., a probabilidade de ser adimplente diminui, em média, 0,14 relativamente ao solicitante com saldo maior que 200.

Tabela 5 Efeitos parciais médios

Em conformidade com a análise da Tabela 5, verificamos que existem algumas variáveis sinalizadas *, cujas interpretações divergem do expectável. Ou seja, quando se esperava um aumento na probabilidade de ser adimplente, a probabilidade diminuiu e quando se esperava uma diminuição, a probabilidade aumentou. Isto pode ocorrer devido a fatores que nem sempre são considerados.

Em suma, podemos dizer que tanto os efeitos parciais das variáveis contínuas como os das variáveis discretas (*dummy*) exercem uma forte influência sobre a probabilidade da variável dependente, com exceção das variáveis já referidas.

3.6 Análise do Poder Preditivo do Modelo

Após a estimação do modelo e interpretação do sinal das variáveis, segue-se a avaliação da sua capacidade preditiva. Deste modo, a classificação dos clientes adimplentes ou inadimplentes baseia-se na probabilidade de adimplência, sendo 0,5 o ponto de corte utilizado, que é o padronizado para o modelo *probit*. A sua estimação é feita do seguinte modo:

$$\tilde{Y} = \begin{cases} 1 & \text{se } \hat{y} \geq 0,5 \\ 0 & \text{se } \hat{y} < 0,5 \end{cases}$$

Sendo que:

$$\hat{P}(\text{credibilidade} = 1|\mathbf{X}) = \hat{y} = \hat{p}$$

$$\widehat{\text{Credibilidade}} = \begin{cases} 1 & \text{se } \hat{p} \geq 0,5 \\ 0 & \text{se } \hat{p} < 0,5 \end{cases}$$

onde \hat{p} é a estimativa da probabilidade condicional de “credibilidade” assumir o valor 1 dada as respostas das variáveis independentes e o 0,5 representa a probabilidade do evento acontecer. Assim, todos solicitantes de crédito com probabilidade estimada de adimplência superior a 0,5 são classificados como adimplentes (pertencentes ao grupo bom crédito) e aqueles com valor da probabilidade de adimplência inferior a 0,5 são classificados como inadimplentes (pertencentes ao grupo mau crédito).

Com objetivo de encontrar bons resultados de previsão, utilizamos mais do que um ponto de corte com valores entre 0,5 e 0,7, sendo este último a proporção de bons clientes na amostra. A Tabela 6 apresenta os resultados gerados pela especificação do poder preditivo do modelo.

Ponto de corte	0,5	0,55	0,6	0,65	0,7
% de clientes inadimplentes	44%	52,33%	58,33%	66,33%	73%
% de clientes adimplentes	91,14%	86,14%	80,86%	75,57%	69,29%
% total de clientes	77%	76%	74,10%	72,80%	70,40%

Tabela 6 Poder preditivo do modelo 2

Os resultados apresentados na Tabela 6, mostram que o ponto de corte que apresenta melhores resultados é o 0,5 (Anexo 5) com um total de 77% dos valores corretamente previstos. De acordo com (Mantovani, Gonçalves, & Gouvêa, 2013) para Picinini, Oliveira e Monteiro (2003) e Selau e Ribeiro (2009), os “modelos de *credit scoring* com taxas de acerto acima de 65% são considerados bons”. Seguindo o pensamento dos

autores, podemos dizer que o nosso modelo não foge à regra, visto que o mesmo está com uma taxa de acerto acima dos 65%.

Entretanto, verificamos que a percentagem de **acertos** do modelo se mostra mais eficiente na classificação dos clientes adimplentes do que na dos clientes inadimplentes. Isto pode acontecer pelo facto de o número de clientes classificados como adimplentes ser maior, ou seja, a amostra é composta por 1000 clientes e 700 são classificados adimplentes e os restantes 300 inadimplentes.

De um modo geral podemos dizer que a implementação do modelo *probit* no desenvolvimento do modelo *credit scoring*, confirma que os resultados gerados pela estimação do modelo são admissíveis. Logo, o modelo estimado pode ser aplicado no processo da tomada de decisão de crédito do banco em análise, contribuindo massiva e positivamente nas estratégias deste processo.

Capítulo 4 Conclusão

Tendo em conta a importância de métodos quantitativos no apoio à tomada de decisão de concessão ou não de crédito, o objetivo deste trabalho foi aplicar o modelo *probit* no desenvolvimento de um modelo de *credit scoring* a partir de dados extraídos da carteira de crédito de um banco alemão.

Na estimação do modelo, consideraram-se 13 variáveis independentes. Sendo a maioria qualitativa foram inseridas outras variáveis no modelo, sob a forma de *dummies*, perfazendo um total de 25 variáveis. Porém, nem todas as variáveis consideradas mostraram ser estatisticamente significativas para explicar o modelo. Assim, as variáveis não significativas, foram retiradas do modelo e trabalhou-se apenas com as variáveis estatisticamente significativas.

O modelo desenvolvido apresentou resultados satisfatórios relativamente ao poder preditivo de classificação dos clientes quanto ao grupo a que pertence (bom crédito ou mau crédito). De facto, considerando o ponto de corte de 0,5 previu-se corretamente 77% do total dos clientes. O modelo mostrou-se também mais eficiente na classificação correta dos clientes adimplentes do que na dos clientes inadimplentes. Tal pode resultar de fatores não observados no modelo como, por exemplo, variáveis não incluídas no modelo que podem explicar a variável dependente, ou pela desigualdade numérica dos grupos que compõem a amostra.

Em suma, o modelo desenvolvido neste trabalho proporciona informações relevantes que podem melhor auxiliar as decisões dos analistas de crédito do banco em análise. Este modelo mostra ser uma ferramenta promissora para análise de crédito, podendo ainda ser aperfeiçoado com a inclusão de outras variáveis relevantes para o processo de decisão.

As informações obtidas do modelo, podem conduzir a um estudo mais detalhado sobre os clientes mais inclinados para inadimplência.

Todavia, é imprescindível que os bancos e outras empresas financeiras façam sempre atualizações nos seus modelos de gestão de risco de crédito ou *credit scoring*, observando não só as variáveis que consideram importantes para o processo de crédito, mas também fatores externos como movimentos cíclicos da economia que afetam direta ou indiretamente o mercado e os solicitantes de crédito.

Sugere-se assim, para pesquisas futuras, o desenvolvimento do modelo com a mesma base de dados aplicando outras técnicas para a análise de crédito ou *credit scoring* tais como análise discriminante, algoritmos genéticos, redes neuronais e regressão logística face ao modelo *probit*.

Referências Bibliográficas

Araújo, E. A., & Carmona, C. U. (2007). Desenvolvimento de Modelos *Credit Scoring* com Abordagem de Regressão Logística para a Gestão da Inadimplência de uma Instituição de Microcrédito.

Base de dados. Disponível em:

<http://www.ipipan.eu/~teisseyp/TEACHING/ZMUM/DANE/kredit.htm> [Acesso em: 02/12/2016]

Bastardo, C. (2011). *Gestão de Activos Financeiros - Back to Basis*. Lisboa: Escolar Editora.

Bessis, J. (2010). *Risk Management in Banking* (3ª ed.). Chichester United Kingdom: John Wiley & Sons Ltd.

Blatt, A. (1999). *Avaliação de Risco e Decisão de Crédito - um enfoque prático*. São Paulo: Nobel.

Caiado, J., & Caiado, A. C. (2009). *Gestão de Instituições Financeiras* (2ª ed.). Lisboa: Edições Silabo.

Carrilho, P. Q. (2009). *O Seu Primerio Milhão - Como poupar e fazer crescer seu dinheiro* (5ª ed.). Alfragide: Lua de Papel.

Casu, B., Girardone, C., & Molyneux, P. (2006). *Introduction to Banking*, England, Prentice Hall Financial Times.

Chaia, A. J. (2003). *Modelo de Gestão do Risco de Crédito e sua Aplicabilidade ao Mercado brasileiro*. Dissertação (Mestrado) em Administração e Contabilidade, Universidade de São Paulo.

Deutsche Bank. (1 de Julho de 2017). *Preçário Deutsche Bank AG*, Sucursal em Portugal. Portugal.

Fama, R., & Santos, J. O. (2007). Avaliação da Aplicabilidade de um Modelo de *Credit Scoring* com variáveis Sistemáticas e Não-Sistemáticas em carteiras de Crédito Bancário Rotativo de pessoas físicas. *Revista de Cont. Fin.* (44), 105-117.

- Mantovani, D. M., Gonçalves, E. B., & Gouvêa, M. A. (2013). Análise de Risco de Crédito com o uso de Regressão Logística. *R. Contemporânea de Contabilidade*, 10 (20), 139-160.
- Morais, J. M. (2011). *Risco de Crédito e o Efeito dos Colaterais*. Estágio (Mestrado) em Económia, Direcção Comercial Empresas Santander Totta, Faculdade de Economia, Coimbra.
- Neto, A. A., & Carmona, C. U. (2004). Modelagem do Risco de Crédito: Um estudo do segmento de pessoas físicas em um banco de varejo. *READ*, 10 (4), 2-23.
- Neves, J. C. (2001). *Princípios de Economia Política* (2ª ed.). São Paulo: Verbo.
- Santos, J. O. (2003). *Análise de Crédito - Empresas e Pessoas Físicas* (2ª ed.). São Paulo: Atlas.
- Silva, P. R. (2011). *Psicologia do Risco de Crédito - Análise da contribuição de variáveis psicológicas em Modelos de Credit Scoring*. Tese (Doutoramento) em Administração, Universidade de São Paulo.
- Vaz, J. L. (1999). *Gestão Bancária*. Lisboa: Editora Universidade Aberta.
- Veloso, G. M. (2016). *Gestão de Risco de Crédito - Abordagem ao rating interno do Banco*. Estágio (Mestrado) em Gestão, Universidade Católica Portuguesa, Porto.
- Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach* (5 ed.). New York: South-Western College Publishers.

Anexo 1 Resumo das variáveis da base de dados

Variável	Descrição	Categorias	Score	Frequência rel. em % para	
				Bons créditos	Maus créditos
AD	Ativos disponíveis	Posse de casa ou terra	4	22.33	12.43
		Contrato de poupança com uma sociedade de construção / Seguro de vida	3	34.00	32.86
		Carro / Outros	2	23.67	23.00
		Sem ativos disponíveis	1	20.00	31.71
DC	Duração do crédito em meses (categorizada)	<=6	10	3.00	10.43
		6 < ... <= 12	9	22.33	30.00
		12 < ... <= 18	8	18.67	18.71
		18 < ... <= 24	7	22.00	22.57
		24 < ... <= 30	6	6.33	5.43
		30 < ... <= 36	5	12.67	6.86
		36 < ... <= 42	4	1.67	1.71
		42 < ... <= 48	3	10.67	3.14
		48 < ... <= 54	2	0.33	0.14
		> 54	1	2.33	1.00
EC/S	Estado civil / Sexo	Masculino: divorciado / vivendo separado	1	6.67	4.29
		Feminino: divorciado / vivendo separado / casado	2	11.33	10.29
		Masculino: solteiro	2	25.00	18.43
		Homem: casado / viúvo	3	48.67	57.43
		Feminino: solteiro	4	8.33	9.57
IDADE	Anos de idade (categorizado)	0 <= ... <= 25	1	26.67	15.71
		26 <= ... <= 39	2	47.33	52.72
		40 <= ... <= 59	3	21.67	26.14
		60 <= ... <= 64	5	2.33	3.00
		>= 65	4	2.00	2.43
MC	Montante do crédito em DM (categorizado)	<=500	10	1.00	2.14
		500 < ... <= 1000	9	11.33	9.14
		1000 < ... <= 1500	8	17.00	19.86
		1500 < ... <= 2500	7	19.67	24.57
		2500 < ... <= 5000	6	25.00	28.57
		5000 < ... <= 7500	5	11.33	9.71
		7500 < ... <= 10000	4	6.67	3.71
		10000 < ... <= 15000	3	7.00	2.00
		15000 < ... <= 20000	2	1.00	0.29
		> 20000	1	0.00	0.00

Variável	Descrição	Categorias	Score	Frequência rel. em % para	
				Bons créditos	Maus créditos
NCANB	Número de créditos anteriores neste banco (incluindo o que está em execução)	1	1	66.67	61.86
		2 ou 3	2	30.67	34.43
		4 ou 5	3	2.00	3.14
		6 ou mais	4	0.67	0.57
ND	Número de dependentes	0 a 2	2	84.67	84.43
		3 e mais	1	15.33	15.57
OC	Objetivo do Crédito	Carro novo	1	5.67	12.29
		Carro usado	2	19.33	17.57
		Artigos mobiliários	3	20.67	31.14
		Rádio / Televisão	4	1.33	1.14
		Eletrodomésticos	5	2.67	2.00
		Reparações	6	7.33	4.00
		Educação	7	0.00	0.00
		Férias	8	0.33	1.14
		Reeducação	9	11.33	9.00
		Negócios	10	1.67	1.00
		Outros	0	29.67	20.71
OCE	Outros créditos em execução	Em outros bancos	1	19.00	11.71
		Na loja de departamentos ou na casa de ordem de correio	2	6.33	4.00
		Sem mais crédito em execução	3	74.67	84.29
OCUP	Ocupação	Desempregado/não qualificado sem residência permanente	1	2.33	2.14
		Empregado não qualificado com residência permanente	2	18.67	20.57
		Funcionário público com exceção de quadro superior	3	62.00	63.43
		Executivo/por conta própria/quadro superior da função pública	4	17.00	13.86
OD/G	Outros devedores / garantes	Nenhum	1	90.67	90.71
		Co-requerentes	2	6.00	3.29
		Fiador	3	3.33	6.00
SCC	Saldo da conta corrente	Sem saldo ou débito	2	35.00	23.43
		0 <= ... < 200 DM	3	4.67	7.00
		... >= 200 DM ou conta corrente por pelo menos 1 ano	4	15.33	49.71
		Sem movimento na conta	1	45.00	19.86

Variável	Descrição	Categorias	Score	Frequência rel. em % para	
				Bons créditos	Maus créditos
SPCE	Situação perante créditos existentes	Sem créditos anteriores/pagou todos créditos anteriores	2	56.33	51.57
		Créditos anteriores pagos neste banco	4	16.67	34.71
		Sem problemas com créditos atuais neste banco	3	9.33	8.57
		Falhas no pagamento de créditos anteriores	0	8.33	2.14
		Conta corrente problemática/há mais créditos em execução em outros bancos	1	9.33	3.00
T.ESTRANG	Trabalhador estrangeiro	Sim	1	1.33	4.71
		Não	2	98.67	95.29
TA	Tipo de apartamento	Apartamento alugado	2	62.00	75.43
		Apartamento ocupado pelo proprietário	3	14.67	9.14
		Apartamento gratuito	1	23.33	15.57
TAXA	Prestação em % do rendimento disponível	>= 35	1	11.33	14.57
		25 <= ... < 35	2	20.67	24.14
		20 <= ... < 25	3	15.00	16.00
		< 20	4	53.00	45.29
TE	Tempo de emprego	desempregado	1	7.67	5.57
		<= 1 ano	2	23.33	14.57
		1 <= ... < 4anos	3	34.67	33.57
		4 <= ... < 7 anos	4	13.00	19.29
		>= 7 anos	5	21.33	27.00
TEL	Telefone	Não	1	62.33	58.43
		Sim	2	37.67	41.57
TR	Tempo de residência	< 1 ano	1	12.00	13.43
		1 <= ... < 4 anos	2	32.33	30.14
		4 <= ... < 7 anos	3	14.33	15.14
		>= 7 anos	4	41.33	41.29
VPS	Valor de poupança ou <i>stock</i>	< 100,- DM	2	11.33	9.86
		100,- <= ... < 500,- DM	3	3.67	7.43
		500,- <= ... < 1000,- DM	4	2.00	6.00
		>= 1000,- DM	5	10.67	21.57
		Não disponível / sem poupança	1	72.33	55.14

Tabela 7 Frequência relativa das variáveis categóricas segundo o tipo de crédito

Anexo 2 Estatística descritiva

Variável	Média	Máximo	Mínimo	Desvio padrão
AUTO	0,284000	1,000000	0,000000	0,451162
CASA	0,364000	1,000000	0,000000	0,481389
CRED	0,700000	1,000000	0,000000	0,458487
DC	20,90300	72,00000	4,000000	12,05881
DE_0_A_200	0,063000	1,000000	0,000000	0,243085
DESEMP	0,062000	1,000000	0,000000	0,241276
EDUC	0,097000	1,000000	0,000000	0,296106
EMPM4	0,427000	1,000000	0,000000	0,494890
FER	0,009000	1,000000	0,000000	0,094488
FPCA	0,040000	1,000000	0,000000	0,196057
FPEQS	0,630000	1,000000	0,000000	0,483046
IDADE	35,54200	75,00000	19,00000	11,35267
LD	0,047000	1,000000	0,000000	0,211745
MC	3271,248	18424,00	250,0000	2822,752
MCEOB	0,049000	1,000000	0,000000	0,215976
NCM6	0,006000	1,000000	0,000000	0,077266
NDM3	0,845000	1,000000	0,000000	0,362086
NEGO	0,012000	1,000000	0,000000	0,108940
OB	0,139000	1,000000	0,000000	0,346120
PCT	0,154000	1,000000	0,000000	0,361129
PRESTM20_A_25	0,157000	1,000000	0,000000	0,363983
PRESTM25_A_35	0,231000	1,000000	0,000000	0,421683
PRESTM35	0,136000	1,000000	0,000000	0,342960
QSFP	0,148000	1,000000	0,000000	0,355278
SMC	0,274000	1,000000	0,000000	0,446232
SSD	0,269000	1,000000	0,000000	0,443662

Tabela 8 Estatística descritiva

Anexo 3 Correlação da variável dependente com outras variáveis

	CRED
AUTO	0,049360
CASA	0,059867
CRED	1,000000
DC	-0,214927
DE_0_A_200	0,044009
DESEMP	-0,039815
EDUC	-0,036129
EMPM4	0,110732
FER	0,039281
FPCA	-0,144767
FPEQS	0,013559
IDADE	0,091272
LD	-0,050523
MC	-0,154740
MCEOB	-0,134448
NCM6	-0,005651
NDM3	-0,003015
NEGO	-0,028058
OB	-0,096510
PCT	-0,125750
PRESTM20_A_25	0,012596
PRESTM25_A_35	0,037796
PRESTM35	0,043289
QSFP	-0,040559
SMC	-0,258333
SSD	-0,119581

Tabela 9 Correlação

Anexo 4 Teste de significância

Variáveis *fer* e *nego*;

Redundant Variables Test
 Null hypothesis: FER NEG0 are jointly insignificant
 Equation: EQ01
 Specification: CRED C LOG(MC) IDADE DC AUTO CASA EDUC FER
 NEG0 PRESTM35 PRESTM25_A_35 PRESTM20_A_25 OB LD FPCA
 MCEOB NCM6 DESEMP EMPM4 FPEQS QSFP PCT NDM3 SMC SSD
 DE_0_A_200
 Redundant Variables: FER NEG0

	Value	df	Probability
Likelihood ratio	4.945674	2	0.0843

LR test summary:

	Value	df
Restricted LogL	-497.2146	976
Unrestricted LogL	-494.7417	974

Tabela 10 Teste de significância conjunta de *fer* e *nego*

$$H_0: \beta_7 = \beta_8 = 0$$

$$H_1: H_0 \text{ falsa}$$

$$LR = 4,945$$

$$P_{obs} = 0,0843 > 0,05$$

Variáveis *ob* e *ld*;

Redundant Variables Test
 Null hypothesis: OB LD are jointly insignificant
 Equation: EQ01
 Specification: CRED C LOG(MC) IDADE DC AUTO CASA EDUC FER
 NEG0 PRESTM35 PRESTM25_A_35 PRESTM20_A_25 OB LD FPCA
 MCEOB NCM6 DESEMP EMPM4 FPEQS QSFP PCT NDM3 SMC SSD
 DE_0_A_200
 Redundant Variables: OB LD

	Value	df	Probability
Likelihood ratio	2.588179	2	0.2741

LR test summary:

	Value	df
Restricted LogL	-496.0358	976
Unrestricted LogL	-494.7417	974

Tabela 11 Teste de significância conjunta de *ob* e *ld*

$$H_0: \beta_{12} = \beta_{13} = 0$$

$$H_1: H_0 \text{ falsa}$$

$$LR = 2,588$$

$$P_{obs} = 0,274 > 0,05$$

Variáveis *fpeqs* e *qsfp*.

Redundant Variables Test
 Null hypothesis: FPEQS QSFP are jointly insignificant
 Equation: EQ01
 Specification: CRED C LOG(MC) IDADE DC AUTO CASA EDUC FER
 NEG0 PRESTM35 PRESTM25_A_35 PRESTM20_A_25 OB LD FPCA
 MCEOB NCM6 DESEMP EMPM4 FPEQS QSFP PCT NDM3 SMC SSD
 DE_0__A__200
 Redundant Variables: FPEQS QSFP

	Value	df	Probability
Likelihood ratio	0.292281	2	0.8640

LR test summary:

	Value	df
Restricted LogL	-494.8879	976
Unrestricted LogL	-494.7417	974

Tabela 12 Teste de significância conjunta de *fpeqs* e *qsfp*

$$H_0: \beta_{19} = \beta_{20} = 0$$

$$H_1: H_0 \text{ falsa}$$

$$LR = 0,292$$

$$P_{obs} = 0,864 > 0,05$$

Teste de todas variáveis não significativas

Redundant Variables Test
 Null hypothesis: LOG(MC) FER NEG0 PRESTM20_A_25 OB LD NCM6 D...
 Equation: EQ01
 Specification: CRED C LOG(MC) IDADE DC AUTO CASA EDUC FER
 NEG0 PRESTM35 PRESTM25_A_35 PRESTM20_A_25 OB LD FPCA
 MCEOB NCM6 DESEMP EMPM4 FPEQS QSFP PCT NDM3 SMC SSD
 DE_0__A__200
 Redundant Variables: LOG(MC) FER NEG0 PRESTM20_A_25 OB LD
 NCM6 DESEMP FPEQS QSFP NDM3

	Value	df	Probability
Likelihood ratio	8.934375	11	0.6279

LR test summary:

	Value	df
Restricted LogL	-499.2089	985
Unrestricted LogL	-494.7417	974

Tabela 13 Teste de significância conjunta de todas não significativas

$$H_0: \beta_1 = \beta_7 = \beta_8 = \beta_{11} = \beta_{12} = \beta_{13} = \beta_{16} = \beta_{17} = \beta_{19} = \beta_{20} = \beta_{22} = 0$$

$$H_1: H_0 \text{ falsa}$$

$$LR = 8,934$$

$$P_{obs} = 0,6279 > 0,05$$

Atendendo aos resultados dos testes concluímos que a hipótese nula H_0 em todos os casos, não é rejeitada, logo não há evidência de significância conjunta das variáveis testadas.

Anexo 5 Estimação do poder preditivo

Expectation-Prediction Evaluation for Binary Specification

Equation: EQ02

Date: 10/06/17 Time: 17:22

Success cutoff: C = 0.5

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	132	62	194	0	0	0
P(Dep=1)>C	168	638	806	300	700	1000
Total	300	700	1000	300	700	1000
Correct	132	638	770	0	700	700
% Correct	44.00	91.14	77.00	0.00	100.00	70.00
% Incorrect	56.00	8.86	23.00	100.00	0.00	30.00
Total Gain*	44.00	-8.86	7.00			
Percent Gain**	44.00	NA	23.33			

Tabela 14 Estimação do poder preditivo