

INSTITUTO POLITÉCNICO DE LISBOA  
INSTITUTO SUPERIOR DE CONTABILIDADE  
E ADMINISTRAÇÃO DE LISBOA



ISCAL

**Modelo de *Scoring* para o Crédito à Habitação:  
Um estudo de caso em Portugal**

---

Joana Isabel Cabrito Vicente

Lisboa, Julho de 2015



INSTITUTO POLITÉCNICO DE LISBOA  
INSTITUTO SUPERIOR DE CONTABILIDADE  
E ADMINISTRAÇÃO DE LISBOA

## **Modelo de *Scoring* para o Crédito à Habitação: Um estudo de caso em Portugal**

Joana Isabel Cabrito Vicente

Dissertação submetida ao Instituto Superior de Contabilidade e Administração de Lisboa, para cumprimento dos requisitos necessários à obtenção do grau de Mestre, realizada sob a orientação científica do Mestre (Especialista) José Nuno Teixeira de Abreu de Albuquerque Sacadura, professor adjunto na área de Finanças

Constituição do Júri:

Presidente - Doutora Ana Maria de Sotomayor

Arguente – Doutor Joaquim Paulo de Carvalho

Vogal – Especialista (Mestre) José Nuno Sacadura

Lisboa, Julho de 2015

Declaro ser a autora desta dissertação, que constitui um trabalho original e inédito, que nunca foi submetido (no seu todo ou qualquer das suas partes) a outra instituição de ensino superior para obtenção de um grau académico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas. Mais acrescento que tenho consciência de que o plágio – a utilização de elementos alheios sem referência ao seu autor – constitui uma grave falta de ética, que poderá resultar na anulação da presente dissertação.

## **Agradecimentos**

Gostaria de prestar estas palavras com todo o apreço às pessoas que mais me ajudaram nesta fase.

Sem obedecer a nenhuma ordem pré concebida, começo por agradecer ao meu orientador, Mestre (Especialista) José Nuno Teixeira de Abreu de Albuquerque Sacadura, toda a ajuda, todo o tempo que me disponibilizou, pela brevidade de resposta sempre que necessário e por todo o conhecimento que me transmitiu.

De forma especial, um forte bem-haja à minha entidade empregadora, às pessoas que me proporcionaram tudo o que foi necessário para a realização deste trabalho, que deram o ‘Ok’ para que tudo fosse possível, um obrigado à Dra. Marta Paula, ao Dr. Augusto Damas e à Dra. Dulce Silva.

Um obrigado ao meu colega, Dr. Miguel Almeida pela disponibilidade imediata.

Seguidamente, agradecer aos meus amigos e familiares por todo o apoio nesta etapa tão importante da minha vida, e por tudo o que vivemos juntos neste percurso académico.

Um forte agradecimento aos docentes do Instituto Superior de Contabilidade e Administração de Lisboa, pelos conhecimentos adquiridos, pelo companheirismo e pela atenção durante este percurso.

Finalizando, agradeço aos meus pais por me terem dado a hipótese de concretizar mais um projeto na minha vida.

A todos,

**Muito Obrigado!**

## Resumo

A concessão de crédito, essência da atividade de uma instituição financeira, implica incorrer no risco de incumprimento. A atual crise financeira agrava esse risco, obrigando a um esforço de capital dos acionistas, traduzido nos rácios exigidos no novo acordo de Basileia. Desta forma, torna-se relevante a reapreciação dos modelos de concessão de crédito existentes.

Através de um estudo de caso, procuramos entender quais as variáveis mais relevantes no momento da concessão de crédito, mais especificamente, no crédito à habitação. Este tipo de crédito é de extrema importância não só em termos das extensas maturidades, como também, em termos de dimensão na Instituição Financeira em estudo.

Para estudar as variáveis, desenvolvemos um modelo de regressão *Logit*, tendo como base as mais utilizadas na bibliografia.

Hoje em dia, existem instituições que não possuem um modelo de apoio à decisão desenvolvido, como se verifica no caso de estudo, em que o modelo de apoio no crédito à habitação ainda se encontra em fase embrionária. Os analistas, na ‘hora da decisão’, baseiam-se essencialmente na experiência, conjugada com os dados de cada cliente. O ponto inicial da investigação será desvendar esses dados, transportando-os de um perfil para variáveis discriminantes e assim, analisar o impacto que cada indicador terá numa situação de concessão de crédito hipotecário.

Os modelos de *Credit Scoring* têm sido cada vez mais utilizados com o intuito de separar os bons dos maus clientes, os que cumprem dos que não cumprem, e aquando este objetivo cumprido estão criadas as condições para a instituição seguir um sentido impulsionador na criação de sucesso, minimizando o risco de crédito.

Com base numa amostra de créditos à habitação existentes numa Instituição de Crédito de referência no mercado português, as evidências sugerem que a idade dos proponentes, os rendimentos, o rácio financiamento/garantia e a relação com a Instituição são as dimensões mais relevantes no momento da concessão de crédito.

Palavras-Chave: risco de crédito, variáveis, *credit scoring*, crédito hipotecário

## Abstract

The granting of credit, essence of activity of a financial institution, means running the risk of default. The current financial crisis exacerbates this risk, requiring a capital effort of the shareholders, translated in the ratios required by the new Basel Accord. Thus, it is important to review existing lending models.

Through a case study, we try to understand what the most important variables at the time of lending, more specifically, the mortgage loans. This type of credit is of utmost importance not only in terms of extended maturities, but also in terms of size in the Financial Institution under study.

To study the variables have developed a logit regression model, with the most relevant base used in the literature.

Today, there are institutions that do not have a model of decision support developed, as in the case study, in which the model support the credit dwelling is still in its infancy. Analysts, on “decision time”, based mainly on experience, combined with the data for each client. The starting point of the investigation will unravel this data, transporting them from one profile to discriminating variables and thus analyze the impact each indicator will have a mortgage lending situation.

The credit scoring models have been increasingly used in order to separate the good from the bad customers, who meet those who do not comply, and when this purpose fulfilled the conditions are created for the institution following a driving sense in creating success while minimizing credit risk.

Based on a housing credits sample existing in a reference credit institution in the Portuguese market, the evidence suggests that the age of tenderers, income, ratio and loan/security and the relationship with the institution are the most relevant dimensions in time the granting of credit.

**Key-Words:** credit risk; variables; credit scoring; mortgage.

# Índice Geral

Agradecimentos.....	v
Resumo.....	vi
Abstract.....	vii
Índice de Figuras.....	xi
Índice de Tabelas.....	xii
Lista de Abreviaturas.....	xiii
CAPÍTULO I – Introdução.....	1
1.1 Objeto de Estudo.....	1
1.2 Objetivos do Estudo.....	1
1.3 Metodologia do Estudo.....	2
1.4 Estrutura do Trabalho.....	2
CAPÍTULO II – Enquadramento Teórico.....	4
2.1 Sector Bancário e a sua Evolução.....	4
2.1.1 Evolução da atividade bancária para o Crédito à Habitação.....	5
2.1.2 O endividamento Imobiliário em Portugal.....	7
2.1.3 A atualidade do Crédito à Habitação em Portugal.....	9
2.2 O novo Acordo de Capital – Basileia III.....	13
2.2.1 Novas Medidas – Não constantes em Acordos anteriores.....	15
2.3 Crédito à Habitação.....	18
2.3.1 No que consiste a apreciação de um Crédito Hipotecário?.....	18
2.4 Risco de Crédito.....	19
2.5 Modelos de <i>Credit Scoring</i> .....	21
2.5.1 Seleção do Modelo.....	23
2.6 <i>Credit Scoring</i> .....	25
2.6.1 Vantagens e Desvantagens do <i>Scoring</i> .....	29
2.7 Algumas Técnicas Estatísticas utilizadas em <i>Credit Scoring</i> .....	32
2.7.1 Métodos utilizados em <i>Credit Scoring</i> .....	33
2.7.2 Modelos de Regressão.....	33
2.7.3 Regressão Linear e Logística.....	34
2.7.4 Análise Discriminante.....	38
2.7.5 Modelos de Seleção das variáveis.....	39

CAPÍTULO III – Metodologia .....	41
3.1 Análise Descritiva.....	42
3.2 Recolha dos dados.....	43
3.2.1 Validação dos Dados em Excel.....	44
3.3 Variáveis.....	44
3.3.1 Atividades desenvolvidas no processo de financiamento imobiliário .....	47
3.4 Método utilizado no Estudo.....	48
3.4.1 Modelo de Regressão Logística - <i>Logit</i> .....	48
3.4.2 Análise Discriminante.....	49
3.5 Ferramentas utilizadas – SPSS.....	51
CAPÍTULO IV – Estudo Empírico .....	52
4.1 Análise Descritiva da Amostra .....	52
4.2 Análise Descritiva das Variáveis.....	52
4.2.1 Idade ( <i>Age</i> ).....	53
4.2.2 Educação ( <i>Education Level</i> ) .....	53
4.2.3 Relação com o Banco ( <i>Relashion Bank</i> ).....	54
4.2.4 Estado Civil ( <i>Divorce</i> ).....	55
4.2.5 Outros empréstimos ( <i>Other Loans</i> ).....	55
4.2.6 Naturalidade do proponente ( <i>Bigcity</i> ).....	56
4.2.7 Regime Laboral ( <i>Labor Bond</i> ) .....	56
4.2.8 Número de Dependentes ( <i>Nr. Dependents</i> ).....	57
4.2.9 Atividade Profissional ( <i>Job</i> ).....	57
4.2.10 Rácio Financiamento/Garantia ( <i>Loan-To-Value</i> ) .....	58
4.2.11 Taxa de Esforço ( <i>Loan-To-Income</i> ) .....	58
4.2.12 Rendimento Anual Líquido Declarado ( <i>Income</i> ).....	59
4.2.13 Valor do Empréstimo ( <i>Loan Amount</i> ) .....	60
4.2.14 Despesas ( <i>Outgoings</i> ).....	60
4.2.15 Propósito do Empréstimo ( <i>Purpose</i> ) .....	61
4.2.16 Duração do Empréstimo ( <i>Loan Duration</i> ).....	61
4.2.17 Prestação do Empréstimo ( <i>Loan Provision</i> ) .....	62
4.2.18 Garantia ( <i>Guarantee</i> ) .....	62
4.3 Modelo de Regressão .....	63
4.3.1 O Procedimento no SPSS.....	64
4.3.2 Diagnóstico de Multicolinearidade .....	68

CAPÍTULO V – Conclusão.....	70
5.1 Limitações e Sugestões para Trabalhos Futuros.....	70
Referências Bibliográficas.....	72

## Índice de Figuras

Figura 2.1 - Preços da habitação em termos reais .....	6
Figura 2.2 - Dívida Total dos Particulares   Comparação na UE .....	7
Figura 2.3 - Comparação do Rácio do Crédito em Risco por segmento e rácio de cobertura, respetivamente.....	9
Figura 2.4 - Empréstimos do Setor Financeiro Residente a Particulares/ Taxa de Variação Anual.....	10
Figura 2.5 - Dívida dos Particulares   Posições em Fim de Período .....	12
Figura 2.6 - Principais eventos de risco desde 1999 .....	13
Figura 2.7 - Sistema de <i>Score</i> de Crédito.....	29
Figura 2.8 - <i>Terminology for Simple Regression</i> .....	35

## Índice de Tabelas

Tabela 4.1 - Classificação Total da Amostra .....	52
Tabela 4.2 - Descrição da Variável – Idade ( <i>AGE</i> ) .....	53
Tabela 4.3 - Descrição da variável – Nível de Educação ( <i>EDUCATION LEVEL</i> ).....	54
Tabela 4.4 - Descrição da Variável – Relação Bancária ( <i>RELASHION BANK</i> ) .....	54
Tabela 4.5 - Descrição da Variável – Estado Civil ( <i>DIVORCE</i> ).....	55
Tabela 4.6 - Descrição da Variável - Outros empréstimos ( <i>OTHER LOANS</i> ).....	55
Tabela 4.7 - Descrição da Variável – Naturalidade ( <i>BIGCITY</i> ) .....	56
Tabela 4.8 - Descrição da Variável – Regime Laboral ( <i>LABOR BOND</i> ).....	56
Tabela 4.9 - Descrição da Variável – Nr. de dependentes ( <i>Nr. OF DEPENDENTS</i> ) .....	57
Tabela 4.10 - Descrição da Variável - Rácio F/G ( <i>LOAN TO VALUE</i> ).....	58
Tabela 4.11 - Descrição da Variável - Taxa de esforço ( <i>LOAN TO INCOME</i> ).....	58
Tabela 4.12 - Descrição da Variável – Valor do Empréstimo ( <i>LOAN AMOUNT</i> ) .....	60
Tabela 4.13 - Descrição da Variável – Despesas ( <i>OUTGOINGS</i> ) .....	60
Tabela 4.14 - Descrição da Variável – Propósito do Empréstimo ( <i>PURPOSE</i> ).....	61
Tabela 4.15 - Descrição da Variável – Duração do Empréstimo ( <i>LOAN DURATION</i> )...	61
Tabela 4.16 - Descrição da Variável – Prestação do empréstimo ( <i>LOAN PROVISION</i> ) ..	62
Tabela 4.17 - Descrição da Variável – Garantia ( <i>GUARANTEE</i> ).....	62
Tabela 4.18 - <i>Case Processing Summary</i> .....	64
Tabela 4.19 - <i>Dependent Variable Encoding</i> .....	64
Tabela 4.20 - Tabela de Classificação a,b .....	65
Tabela 4.21 - <i>Omnibus Tests of Models Coefficients</i> .....	65
Tabela 4.22 - <i>Model Summary</i> .....	65
Tabela 4.23 - <i>Table Classification a</i> .....	66
Tabela 4.24 - <i>Variables in the Equation</i> .....	67
Tabela 4.25 - Coeficientes a.....	68

## Lista de Abreviaturas

APEMIP: Associação dos Profissionais e Empresas de Mediação Imobiliária de Portugal  
ATM: Automated Teller Machine  
BCBS: Basel Committee on Banking Supervision  
BES: Banco Espírito Santo  
BFB: Banco FONSECAS & Burney  
BFE: Banco Fomento Exterior  
BdP: Banco de Portugal  
BIS: Bank for International Settlements  
BPA: Banco Privado Atlântico  
BPSM: Banco Pinto e Sotto Mayor  
BTA: Banco Totta e Açores  
CERS: Conselho Europeu do Risco Sistémico  
CMVM: Comissão de Mercado de Valores Mobiliários  
CPP: Crédito Predial Português  
CRC: Central de Responsabilidades de Crédito  
DARC: Departamento de Análise de Risco de Crédito  
EBA: European Banking Authority  
IMI: Imposto Municipal sobre Imóveis  
INE: Instituto Nacional de Estatística  
IRS: Imposto de Rendimento a Singulares  
ISFF: Inquérito à Situação Financeira das Famílias  
MAR: Modelo de Avaliação de Riscos  
OCDE: Organização para a Cooperação e Desenvolvimento Económico  
PIB: Produto Interno Bruto  
UBP: União dos Bancos Portugueses

# CAPÍTULO I – Introdução

## 1.1 Objeto de Estudo

A presente investigação incide sobre os fatores e/ou variáveis que determinam o incumprimento no crédito bancário, especificamente, no crédito à habitação.

Passa pela estruturação de um modelo de *Credit Scoring* que transmita a pontuação ou *score* dada a cada cliente/proposta no crédito hipotecário. Serão pois analisadas variável a variável, não só sociodemográficas, como económicas (inerentes à operação) com o objetivo de juntar os indicadores ideais que darão suporte ao modelo.

## 1.2 Objetivos do Estudo

As relações financeiras são o objeto principal da atividade bancária, o que desencadeia um dos principais problemas enfrentados pelas instituições, a gestão do risco. Durante a atividade surgem riscos por contrapartida das operações realizadas entre ativos e passivos, ou seja, o banco concede créditos em contrapartida de uma expectativa de recebimento futuro. Nos últimos anos, a concessão de crédito tem atingido níveis cada vez maiores por parte das instituições financeiras. A necessidade em captar clientes, por força de uma concorrência competitiva, tem contribuído para que sejam processadas inúmeras propostas de crédito sem qualquer nível de viabilidade. Cabe ao analista de risco perceber se os mutuários possuem uma boa capacidade creditícia ou, se por outra, não têm perfil de aforradores conduzindo a situações de incumprimento/*default*<sup>1</sup>.

O crédito em geral engloba vários tipos de empréstimos, no entanto, para o estudo a realizar o foco principal será o crédito à habitação. O objetivo será entender quais os fatores que levam ao incumprimento neste tipo de financiamento, ou seja, o que contribui para o não pagamento das responsabilidades pecuniárias por parte dos mutuários. Neste sentido, surge a necessidade de um estudo aprofundado sobre a contratação deste tipo de crédito, uma vez que será para muitos particulares o maior financiamento que fazem ao longo da vida e, que de igual forma, se estende durante todo esse período, transformando-se num compromisso financeiro de grande responsabilidade e vínculo contratual. Para a execução deste objetivo, terá de se conhecer o perfil de cada cliente, de forma a detetar quais as características comuns que transportam dificuldade nos pagamentos, ou seja, alto risco. Todas essas características

---

<sup>1</sup> *Default* é um estrangeirismo utilizado para traduzir o termo ‘incumprimento’ face a responsabilidades financeiras.

dizem respeito a um conjunto de variáveis que serão estudadas para, posteriormente, gerarem o ponto de partida na criação de um modelo.

Deste modo, o modelo a ser realizado para auxílio dos analistas de crédito será o *Credit Scoring* através da técnica de Regressão Logística. Para a análise do crédito à habitação, torna-se fulcral a existência de ferramentas de apoio à decisão de modo a eliminar a incerteza e, por conseguinte, o risco.

Tendo presente os problemas associados ao incumprimento, e tendo como base um modelo de ajuda na decisão de concessão de crédito, com os dados reais do perfil de um conjunto de clientes, será construído o modelo de pontuação (*scores*) a clientes particulares.

No entanto, há outros fatores que influenciam a recusa ou a concessão de crédito para além de características pessoais ou sociais. Serão indicadas as variáveis de origem económica e que estão internamente ligadas à proposta de crédito, sendo a taxa de juro, rácio financiamento/garantia, a taxa de esforço, prazo do empréstimo, prestação, entre outras. Todos estes indicadores terão de ser conjugados com as variáveis sociais de modo a perceber qual a relação que têm entre si.

### **1.3 Metodologia do Estudo**

Será realizado um estudo que determinará a existência ou não de incumprimento/*default* consoante as variáveis sejam sociodemográficas ou económicas. O estudo incidirá sobre o crédito na aquisição de bens imóveis dentro de uma instituição bancária, a qual disponibilizou o acesso a todos os dados necessários para a execução do estudo. A investigação basear-se-á nos indicadores quantitativos e qualitativos procedentes na amostra, ou seja, no universo de créditos vivos, versus, créditos em incumprimento.

Em complemento à necessidade de justificação de alguns dados de base do estudo contou-se com o auxílio de entidades como o Instituto Nacional de Estatística (INE), o Banco de Portugal (BP) e Comissão do Mercado de Valores Mobiliários (CMVM).

Tecnicamente, recorreu-se a várias linhas gerais sendo: 1) Recolha, e análise dos dados; 2) construção da base de dados; 3) Técnicas para a seleção das variáveis que melhor se ajustam ao modelo; 4) Aplicação da Regressão Logística para a obtenção do modelo; 5) análise do modelo final com as variáveis selecionadas.

### **1.4 Estrutura do Trabalho**

A presente dissertação está descrita em cinco capítulos. Iniciando pelo **Capítulo I – Introdução** onde advém o objeto do estudo, os objetivos e a metodologia de trabalho. De seguida, o **Capítulo II - Enquadramento Teórico** é apresentada a revisão da literatura relativa ao

*Credit Scoring*, nomeadamente, os vários modelos existentes de apoio à decisão, as vantagens da utilização deste modelo, e técnicas para a construção. **Segue o Capítulo III – Metodologia**, onde são descritos os métodos, passos e conceitos utilizados na análise prática do trabalho.

De seguida o **Capítulo IV – Estudo Empírico**, onde se insere a análise dos dados/amostra; definição das variáveis; execução dos vários testes através do Modelo de Regressão Logística; estimação do modelo de *Credit Scoring*; explicação dos testes ao modelo final; e por fim, apresentação do programa que permitiu com que o trabalho fosse estudado da forma mais fiável, *Software Statistical Package for the Social Science* (SPSS) e das ferramentas que com ele foram utilizadas.

Finalizando, fechou-se o estudo de caso com o **Capítulo V – Conclusões** onde é relatada a súmula do trabalho que foi desenvolvido, e apresentam-se ainda várias sugestões e limitações para o desenvolvimento de trabalhos futuros.

## CAPÍTULO II – Enquadramento Teórico

### 2.1 Sector Bancário e a sua Evolução

Em meados dos anos 70, o mercado bancário estava condicionado, derivado da intervenção do Estado no sector. Para além deste condicionamento também existiam fortes barreiras que impossibilitavam o aparecimento de instituições privadas. As instituições financeiras viam a sua atividade limitada através, da existência de limites de crédito, da especialização existente nas instituições (ausência da banca universal) e da instrumentalização dos bancos feita pelo Estado no financiamento a empresas públicas.

Já na segunda metade dos anos 80, começou a inverter-se a tendência de queda do sector que se verificava até então, uma vez que, com a entrada de Portugal na Comunidade Europeia, começaram a surgir reformas no sistema financeiro, iniciando-se o processo de privatizações e modernização do sector. Iniciou-se o processo de reprivatização dos bancos: Banco Totta e Açores (BTA) em 1989, Banco Privado Atlântico (BPA) em 1990, Banco Espírito Santo (BES) em 1991, Banco FONSECAS & Burney (BFB) em 1991, Crédito Predial Português (CPP) em 1992, União dos Bancos Portugueses (UBP) em 1993, Banco Pinto e Sotto Mayor (BPSM) em 1994 e Banco Fomento Exterior (BFE) em 1996, e foram criadas condições para o aparecimento de novos bancos privados, o que implicou um aumento imediato na concorrência e na racionalidade dos investimentos feitos pelos bancos.

A conjuntura que o país começou a viver após a adesão à Comunidade Europeia trouxe uma forte dinâmica de crescimento ao sector de que são exemplo, a expansão dos bancos no território nacional (abertura de agências), as fusões e aquisições e a modernização tecnológica das instituições. Este crescimento foi impulsionado a partir da década de 90, com a plena liberalização de movimentos de capitais, que permitiu às instituições financeiras portuguesas estabelecer filiais e prestar serviços bancários em qualquer país da então Comunidade Económica Europeia (CEE).

Os bancos portugueses passaram a enfrentar uma concorrência transnacional, o que implicou uma necessidade de ganho de escala, quer por via da internalização dos bancos, quer por via da implementação de modelos de crescimento acelerado.

Após esta fase de reorganização do setor pós período de privatizações, concentrações e crescimento dos principais *players* bancários em Portugal, verifica-se a partir de 2007 mais uma mudança substancial no mercado, com a reestruturação e modernização dos principais ban-

cos, dos quais são exemplo o *rebranding*<sup>2</sup> e a aposta nas tecnologias de informação (modernização e democratização da utilização de sites como meios de transação, a proliferação de meios de transação automáticos como as ATM's).

Com a nova conjuntura recessiva que afetou o país e o sistema financeiro como um todo, os bancos tiveram que implementar novas estratégias de “sobrevivência” como a redução do seu GAP comercial (rácio entre crédito e recursos), novas formas de incremento da sua margem financeira, contenção do risco de crédito e a implementação de planos de recapitalização, que em alguns casos significou a entrada do Estado no capital dos bancos.

### **2.1.1 Evolução da atividade bancária para o Crédito à Habitação**

Nos últimos 40 anos, Portugal foi alvo de uma transformação a vários níveis no sector imobiliário, no âmbito da não previsão dos diferentes métodos de acesso ao financiamento para habitação. Durante este período, existiu um elevado número de famílias a viver em bairros habitacionais sem quaisquer condições de vivência. No entanto, com o passar dos anos, essa realidade sofreu algumas alterações e, atualmente, essas famílias tornaram-se proprietárias de habitação própria. O desenvolvimento do sector imobiliário e, por conseguinte, da construção, foi tão substancial que existe, hoje, uma enorme oferta de imóveis comparativamente aos particulares residentes.

Com o aumento do volume de imóveis para venda, e aumento do desemprego devido à crise que se atravessava, aumentaram também os níveis de endividamento das famílias traduzindo-se num incremento dos níveis de incumprimento.

Os anos 90, foram um marco para a expansão no crédito a particulares em Portugal, ao lado de uma liberalização e privatização do sector financeiro. Nesta data, foram extintas as taxas de juro administrativas e limites de crédito impostos, logo, estavam assim reunidas as condições para um aumento da concessão de crédito.

Há que salientar os dois períodos históricos de um elevado crédito habitacional, sendo um posterior à revolução do 25 de Abril [1974 - 1979] e um segundo de [1994 – 2002], que atingiu cerca de 2% do total da despesa pública.

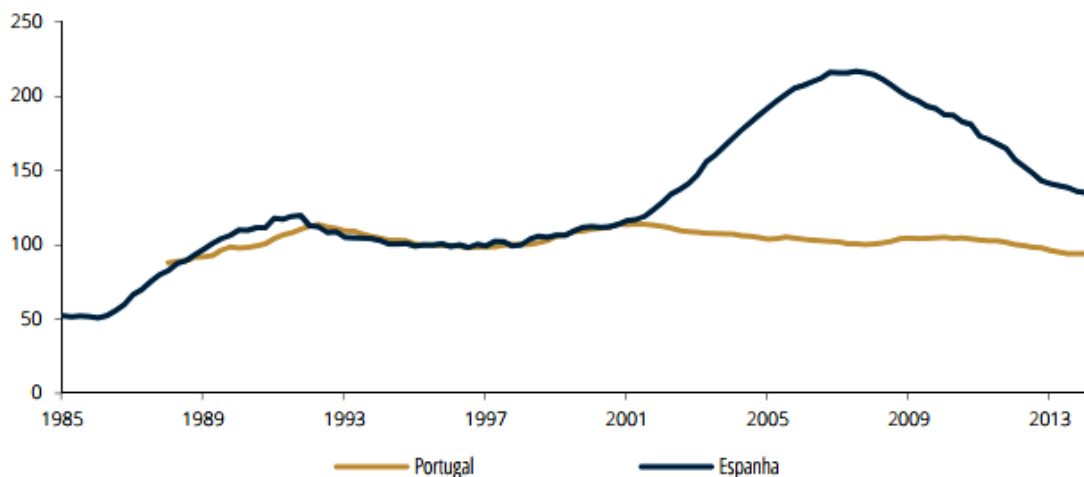
Segundo Santos, Teles e Serra (2014), a título de exemplo, entre 1995 e 2013, o valor dos ativos financeiros, em percentagem do PIB, aumentou cerca de 255 pontos percentuais, representando, em 2013, cerca de 700% do PIB português, embora abaixo de muitos países da zona do Euro.

---

<sup>2</sup> Estratégia de marketing no qual a entidade decide alterar a sua denominação, logotipo ou outros elementos identificativos.

A economia portuguesa teve um forte impacto na concessão de crédito habitação, sendo que surgiram inúmeras políticas sobre a procura de habitação própria e sobre o recurso ao crédito, o qual se tornou de fácil acesso.

**Figura 2.1** - Preços da habitação em termos reais



**Fonte:** OCDE e Banco de Portugal (Boletim Económico/Nov., 2014:41)

Conforme Representado pela figura 2.1,

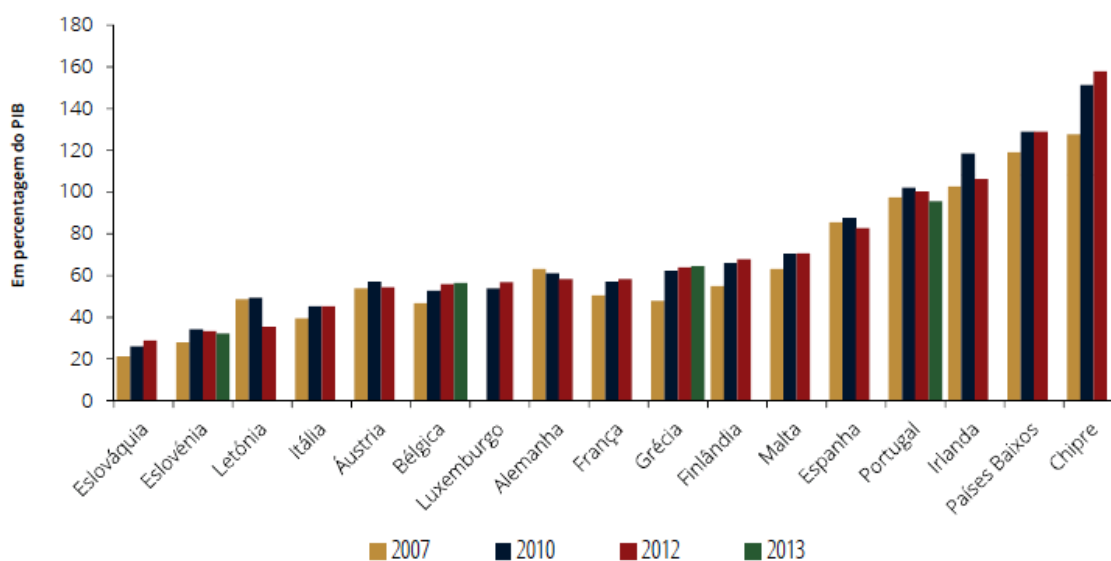
[P]ortugal e Espanha aderiram à UE em 1986 e fizeram simultaneamente o processo de convergência para a moeda única em 1999. É neste contexto, durante duas décadas e até ao eclodir da crise financeira em 2007, que os preços da habitação em termos reais cresceram em média menos de 1 por cento ao ano em Portugal e acima de 6 por cento em Espanha. A evolução em Espanha foi bastante heterogénea quando comparada com Portugal. Entre 1985 e 1998 os preços da habitação aumentaram cerca de 1 por cento ao ano em Portugal e 5 por cento em Espanha. No período entre 1999 e 2006 os preços da habitação registaram um crescimento anual nulo em Portugal, tendo pelo contrário aumentado quase 10 por cento ao ano em Espanha.

Contudo, a partir de 2007 os preços da habitação caíram cerca de 1 por cento ao ano no caso português e 6 por cento no caso espanhol. (BdP-Boletim Económico, 2014:41)

### 2.1.2 O endividamento Imobiliário em Portugal

A necessidade de cobrir bens essenciais e necessidades financeiras das famílias faz com que o recurso ao crédito seja uma constante. A maior parte do endividamento está afeto ao crédito a particulares, do qual faz parte o crédito à habitação, sendo dos financiamentos de maior valor contraídos por um particular ao longo da sua vida, como se pode verificar pela análise da figura 2.2, representada abaixo.

**Figura 2.2 - Dívida Total dos Particulares | Comparação na UE**



**Fonte:** Eurostat (Relatório de Estabilidade Financeira/Nov., BdP 2014:19)

Fatores como a privatização da banca e o não controlo monetário tornaram-se o ponto de partida para o baixo endividamento e para o marco histórico de taxas de juro reduzidas, o que fez com que o crédito à habitação fosse facilmente acessível. Assim haveria uma maior dispersão do risco (dada a diversidade de clientes), a possibilidade uma maior diversificação das taxas de juro (em virtude do seu menor impacto nas prestações, decorrente dos elevados montantes financiados) e a maior maturidade dos empréstimos (Lobo, 1998 *apud* Santos et al., 2014)<sup>3</sup>.

Segundo Santos et al. (2014) ao contrário do que aconteceu nos EUA, em que a concessão do crédito foi atribuída a classes sociais mais baixas e com maior risco de incumprimento, o

<sup>3</sup> Lobo, F. (1998) Crédito ao consumo e restrições de liquidez: uma aplicação à economia portuguesa. Dissertação de Mestrado, Coimbra, Faculdade de Economia, Universidade de Coimbra.

conhecido *Subprime*<sup>4</sup>, o mesmo não aconteceu em Portugal, tendo sido concedido a famílias de maiores rendimentos e, com maiores garantias de solvência financeira. Segundo informações recolhidas pelo Inquérito à Situação Financeira das Famílias (ISFF) efetuada pelo Banco de Portugal (BdP) e Instituto Nacional de Estatística (INE), em 2010 apenas 37,7% das famílias portuguesas estavam endividadas, sendo que 24,5% com empréstimos à habitação e o restante em empréstimos não garantidos por imóveis. O ISFF indica ainda, que quanto maior o rendimento das famílias, maior o aumento da dívida.

O mercado do crédito imobiliário é particularmente elevado em indivíduos com idades entre os 35 e 44 anos, nível de escolaridade superior e trabalhador por conta de outrem com contrato sem termo certo (Costa e Farinha, 2012 *apud* Santos, et al. 2014)<sup>5</sup>.

Segundo os mesmos autores, o crédito não hipotecário, que representa cerca de 8% do total do crédito concedido, tem uma maior incidência em famílias de menor riqueza e rendimento e em indivíduos com idades inferiores a 35 anos, desempregados ou trabalhadores por conta de outrem.

De acordo com o descrito, depreende-se assim que as dívidas hipotecárias são proferidas por famílias cujo rendimentos são mais avultados e com situação profissional estável, contrariamente às dívidas não hipotecárias que indicam uma situação financeira frágil, sendo por vezes o caminho mais fácil para suprir dificuldades de médio prazo (Frade et al., 2008 *apud* Santos et al, 2014)<sup>6</sup>.

Ter-se-á verificado que até emergir a crise económica, o endividamento não mostrava grandes preocupações, foi no entanto após o aparecimento do mesmo que se fez sentir níveis de incumprimento bastante elevados. Com a crise, o endividamento das famílias portuguesas aumentou. Com rendimentos baixos, despontou o crédito malparado nas carteiras dos bancos. A taxa de incumprimento do crédito ao consumo disparou, de 6,7% em 2009 para 12,7% em 2013. Ainda assim, com o pressentimento de que o crédito habitação fosse disparar, a verdade é que se manteve bastante conservador, passando de 1,6% para 2,4% no mesmo período. Na figura 2.3 pode-se visualizar o rácio de crédito em risco por segmento vs. cobertura.

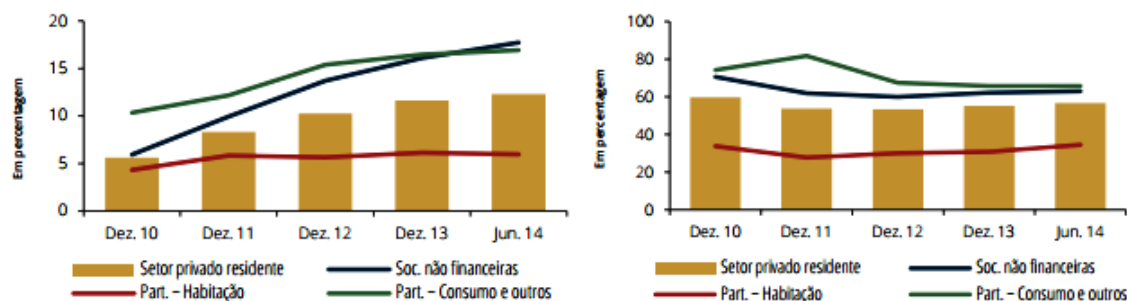
---

<sup>4</sup> “É um crédito à habitação de alto risco, onde determinada população com rendimentos mais baixos acede ao crédito. A única garantia exigida neste tipo de empréstimos é o imóvel. Este segmento do CH é exclusivo nos EUA, não havendo na Europa algo semelhante. (Alves, 2008 *apud* Santos *et al.*, 2014)

<sup>5</sup> Costa, S. e Farinha, L. (2012) “O endividamento das famílias: uma análise microeconómica com base nos resultados do inquérito à situação financeira das famílias” in Banco de Portugal (ed.), Relatório de Estabilidade Financeira Maio 2012, Lisboa, pp.137-163

<sup>6</sup> Frade, C., Lopes, C., Jesus, F. e Ferreira, T. (2008), *Um perfil dos sobre endividados em Portugal*. Coimbra, Centro de Estudos Sociais.

**Figura 2.3** - Comparação do Rácio do Crédito em Risco por segmento e rácio de cobertura, respetivamente



**Fonte:** Banco de Portugal (Dados obtidos a partir da Instrução n.º 22/2011 do BdP)

Em sintonia com o que foi dito anteriormente, verifica-se assim que o crédito à habitação, tendo como garantia o valor do imóvel, e naturalmente associado a taxas de juro mais reduzidas, apresenta uma maior facilidade na gestão do crédito concentrando-se, em escalões de maior rendimento. Nos anos 90, registaram-se níveis de construção muito elevados trazendo ao mercado uma taxa de posse de bens imobiliários em grande escala. Dado esse facto, o aumento da procura também se verificou de forma acentuada, resultando assim numa subida forte dos preços das habitações, principalmente no ano 2000 (ibid).

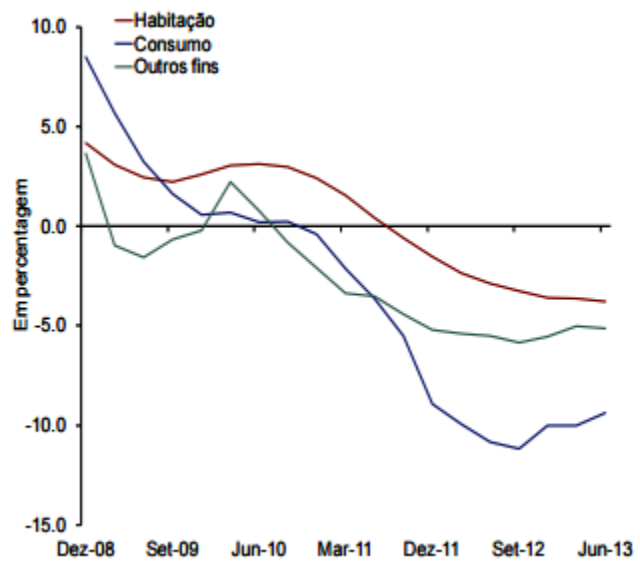
### 2.1.3 A atualidade do Crédito à Habitação em Portugal

Num contexto de quebra do rendimento disponível das famílias e de elevada taxa de desemprego, o rácio de incumprimento continuou a aumentar no crédito ao consumo e a outros fins, mas permanece contido no crédito habitação.

A redução do rendimento disponível das famílias e de elevada taxa de desemprego continuaram a induzir uma significativa materialização do risco de crédito no segmento dos empréstimos bancários a particulares para consumo e outros fins, cujo rácio de crédito de cobrança duvidosa se mantém numa trajetória ascendente desde 2008. Em contrapartida, o rácio de crédito de cobrança duvidosa nos empréstimos para aquisição de habitação verifica uma evolução mais moderada e níveis contidos. Este resultado deve-se ao facto de estes empréstimos respeitarem principalmente à aquisição de habitação permanente, ativo que os garante, pelo que o incumprimento é tendencialmente mais reduzido.

(BdP-Situação patrimonial dos setores não financeiros, 2013:21)

**Figura 2.4** - Empréstimos do Setor Financeiro Residente a Particulares/ Taxa de Variação Anual



**Fonte:** Banco de Portugal (Situação Patrimonial dos setores não financeiros, 2013:21)

No entanto, existem instituições que apostam fortemente no crédito hipotecário, segundo uma notícia no Jornal Público (2014):

O Banco BIC lançou recentemente uma campanha televisiva direcionada para a compra de habitação. Ao contrário dos seus pares, o Banco BIC não tem uma carteira de crédito à habitação, com *spreads* (margem comercial) baixíssimos, a pesar-lhe no balanço, e isso explica o tiro de partida num segmento que estava literalmente parado desde a crise financeira de 2008.

Apesar de não ser assumido objetivamente, a campanha do BIC tenta atrair um público jovem, seguindo a velha estratégia de fidelizar clientes através do crédito à habitação. O *spread* de base da campanha é de 3%, que é somada à taxa Euribor a 3 meses, prova, no entanto, que a entrada na corrida do crédito à habitação pelo banco que comprou o BPN está a ser feita com cautela.

Após esta intervenção, Ricardo Guimarães (Diretor da Confidencial Imobiliário – Revista especializada no setor) e também o Presidente da APEMIP (associação de empresas de me-

dição imobiliária) defendem que os restantes bancos, quando estiverem as condições reunidas, regressarão a este segmento de mercado, os mesmos defendem ainda que dentro de dois anos, os *spreads* estarão entre 1% e 1,5%.

De acordo com a mesma notícia, o Jornal Público refere ainda que

[e]ste movimento já começou. Vários bancos nacionais, incluindo a Caixa Geral de Depósitos, já baixaram o *spread* mínimo de 3,5% para 2,5%. A margem máxima, utilizado para travar operações de alto risco, ainda está entre os 6% e 7%.”

O último inquérito realizado pelo Banco de Portugal aos cinco maiores bancos a operar no país revela que “não obstante o nível do *spread* praticado, os bancos reportaram uma maior diferenciação dos clientes de acordo com o perfil de risco, tanto no crédito à habitação como no crédito ao consumo.

Defende-se que a atividade do segmento de crédito à habitação, é de extrema importância para o setor, pelo número de imóveis em carteira e pelo interesse em renovar contratos, conseguindo-se com troca de casa, e acabar de forma gradual com os *spreads* de 0,25% e 0,30%, percentagens que transportam prejuízos.

A grande aposta da banca embate sobre a população jovem, pela fidelização do cliente e pela quantidade de produtos que poderá subscrever com a concessão de crédito habitação, produzindo uma maior rentabilidade para o banco. No entanto, atualmente este processo complica-se devido ao período anterior à crise.

Na continuidade da notícia pelo Público, “A perceção do risco face ao emprego”, nas palavras de Ricardo Guimarães, afasta os jovens do projeto de comprar casa. Para o diretor da Confidencial Imobiliário, responsável pelo índice com o mesmo nome, que mede a evolução do sector, “a geração pós-25 de Abril revela um certo desapego em relação à propriedade” e “hoje existe um mercado de arrendamento”, ao contrário do que acontecia há quatro anos atrás.

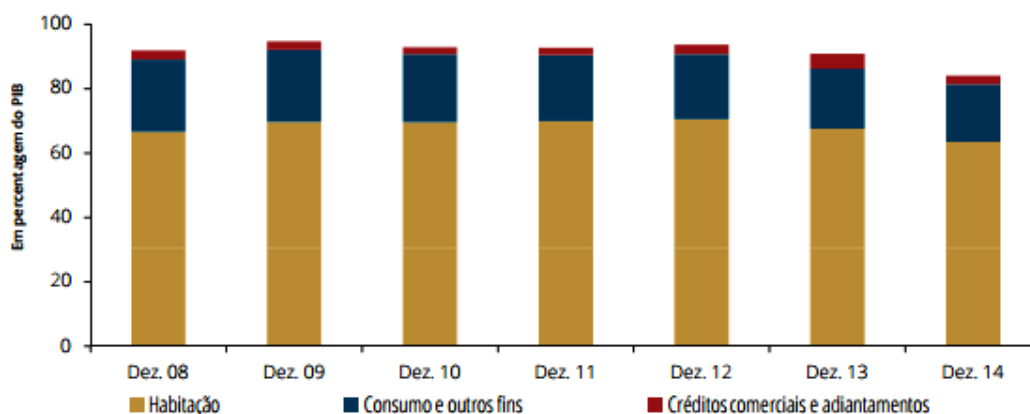
“As rendas continuam altas”, defende o presidente da APEMIP, acrescentado que elas só vão baixar quando o enquadramento fiscal for alterado. “Não se pode penalizar a propriedade [IMI] e o rendimento das rendas [taxa liberatória ou englobamento em IRS]”.

Referem ainda que não há números concretos sobre o mercado de arrendamento, apenas informações sobre algumas imobiliárias defendendo o crescimento desse tipo negócio, representando cerca de 50% da atividade. Os pedidos de visitas cresceram entre

25% e 30% no 1º trimestre de 2014, no entanto, no que diz respeito a vendas efetivas terão crescido apenas 0,5%.

O crédito imobiliário está em mínimos históricos. Em Fevereiro último foi concedido 161 milhões de euros de créditos à habitação (novos contratos). Em Fevereiro de 2008, o valor concedido ascendeu a 1.223 milhões de euros. O acumulado de crédito à habitação nos últimos três anos diminuiu mais de sete mil milhões de euros, de 111.511 milhões para 104.269 milhões. A queda das taxas de juros Euribor, a que estão associados os empréstimos, fez acelerar a componente da amortização e ajudam, juntamente com a travagem nas novas concessões, a explicar a queda do *stock* de crédito concedido. (Jornal Público, 2014)

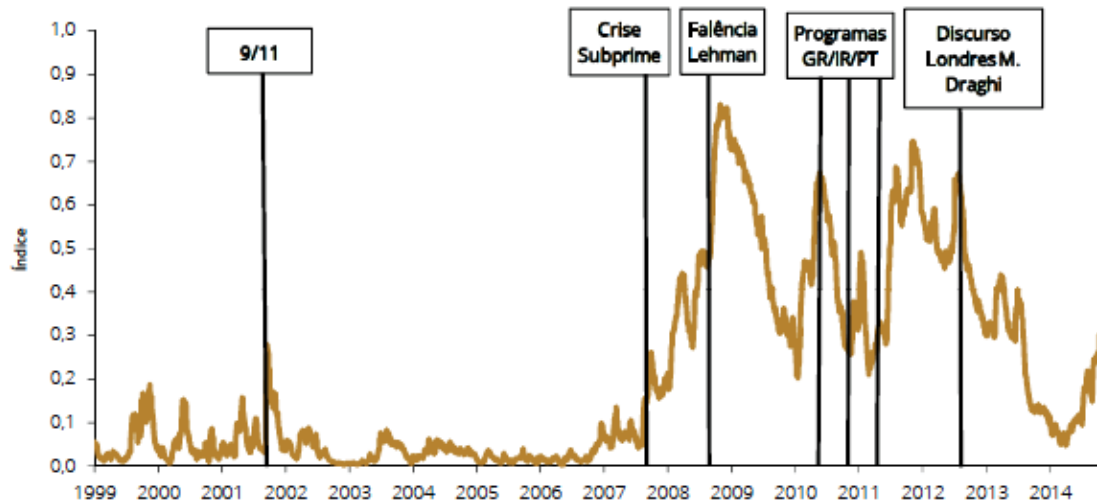
**Figura 2.5 - Dívida dos Particulares | Posições em Fim de Período**



**Fonte:** Banco de Portugal (Relatório de Estabilidade Financeira/Maio, 2015:18)

Desde setembro de 2008, quando a Euribor a seis meses (principal indexante dos contratos de crédito à habitação, que constituem a maior parcela da dívida dos particulares) registou um valor máximo, a prestação de crédito à habitação tem vindo a reduzir-se. Tal facto permitiu atenuar os efeitos da queda do rendimento disponível ocorrida durante o Programa de Assistência Económica e Financeira (PAEF) e, ao mesmo tempo, conter o incumprimento neste segmento do mercado de crédito. No caso de um empréstimo contratado com prazo de 30 anos e spread de um ponto percentual, a prestação mensal ter-se-á reduzido para quase metade entre setembro de 2008 e dezembro de 2014. (Relatório de Estabilidade Financeira/2015, BdP)

**Figura 2.6** - Principais eventos de risco desde 1999



Fonte: Banco de Portugal

O indicador obtido capta de forma adequada os principais eventos de *stress* verificados nos mercados financeiros nacionais.

Destaque para os picos de tensão verificados aquando da falência do Lehman Brothers e no âmbito da crise da dívida soberana que afetou um conjunto de países da área do euro. (Relatório de Estabilidade Financeira, Nov./2014, BdP)

## 2.2 O novo Acordo de Capital – Basileia III

O sistema bancário é composto por uma diversidade de riscos provenientes do decorrer da atividade diária. Pode afirmar-se que, o sistema financeiro engloba um conjunto de entidades, instrumentos e mercados que permitem canalizar as poupanças dos aforradores para as necessidades de financiamento dos mutuários.

Deste modo, cada vez mais são exigidos meios que combatam os riscos inerentes à atividade das instituições havendo uma maior implementação de regras e processos com o objetivo de supervisionar e de tornar transparente todas as relações bancárias, garantindo a continuidade sustentada da instituição.

De acordo com o descrito, houve a necessidade de estabelecer novas regras de capital mínimo, fruto da crise de 2008 e 2009. O Comité de Basileia publicou em 12/Setembro de 2010

o Acordo de Basileia III. Este documento reparte-se em: «*A global regulatory Framework for more resilient banks and banking systems*» (BIS, junho 2011) e «*International Framework for liquidity risk measurement, standards and monitoring*» (BIS, dezembro 2010)

Vários acontecimentos tiveram contribuição no despontar da crise (2008/2009) como o *subprime*, fortes quedas bolsistas, taxas de juro nos EUA, fraca supervisão dos principais bancos, débil gestão do risco. As grandes falências existentes em 2008 por parte do setor financeiro exigiram a reformulação de regras, daí a elaboração de um novo acordo - Basileia III. O novo acordo visa sobretudo acautelar o sistema financeiro, permitir um funcionamento sustentado e minimizar perdas. Com este fim, foram introduzidas regras, regulamentos e requisitos com a finalidade de minimizar o risco (liquidez, crédito, mercado, operacional.).

### **Principais mudanças ao acordo existente:**

No novo acordo de Basileia III existem fortes alterações, as principais são ao nível do capital regulatório, o *Tier I* e o risco de crédito.

Com respeito ao capital regulatório, o novo acordo prevê um reforço da qualidade do mesmo, sendo que ao longo da crise cada vez mais aumentam as imparidades de ativos tóxicos, destruindo assim a base de capital, o que leva à descapitalização dos bancos. Com o aumento da base de capital haverá uma maior margem para acomodar eventuais perdas.

O *Tier I* será composto por capital principal<sup>7</sup> e por capital adicional<sup>8</sup>. Seguidamente, o *Tier II* passa a ser composto por dívidas subordinadas<sup>9</sup>, desde que não exceda em 50% do valor do capital do *Tier I*.

O capital principal, ou *common equity Tier I*

[c]onstitui o capital de melhor qualidade da instituição, em termos de permanência e capacidade de absorção de prejuízos. Salvo no caso de uma eventual capitalização com recurso a investimento público, concretizada através da aquisição de ações pelo Estado com direitos especiais em termos de remuneração, o conceito de Common Equity Tier I corresponde, no início da aplicação das novas regras de Basileia III, ao numerador do rácio *Core Tier I* definido pelo Banco de Portugal. (Comunicado Relativo a aviso do BdP sobre reforço do rácio 'Core Tier 1' das Instituições de Crédito, 2011)

---

<sup>7</sup> Ou *common equity Tier I*: é a soma do capital social do banco, deduzindo dividendos e lucros retidos)

<sup>8</sup> Ou *additional Tier I*: são instrumentos híbridos de capital e dívida

<sup>9</sup> Dívidas subordinadas são dívidas que não são cobertas por garantias reais ou bancárias.

Ainda segundo comunicado do Banco de Portugal, o rácio *Core Tier 1*:

[e]stabelece um nível mínimo de capital que as instituições deveram ter em função dos requisitos de fundos próprios decorrentes dos riscos associados à sua atividade. Como tal, este rácio é apurado através do quociente entre o conjunto de fundos próprios designado de ‘core’ e as posições ponderadas em função do seu risco.

O conjunto de fundos próprios ‘core’ compreende o capital de melhor qualidade da instituição, em termos de permanência e capacidade de absorção de prejuízos, deduzido de eventuais prejuízos e de certos elementos sem valor de realização autónomo (vide lista detalhada de elementos elegíveis em anexo), numa perspetiva de continuidade da atividade de uma instituição.

Todos estes conceitos contribuirão para que o funcionamento das instituições bancárias seja mais rigoroso, numa ótica de maior absorção de perdas e solvabilidade.

Conforme estipulado pelo *Basel Committee on Banking Supervision* (BCBS), no que diz respeito a um segundo ponto de abordagem, os limites mínimos de capital regulatório, os bancos devem manter 4,5% do seu ativo ponderado pelo risco (RWA) do capital principal, 6% do ativo ponderado pelo risco do *Tier I* e 8% do ativo ponderado pelo risco do total do seu capital regulatório. O conjunto destes limites direciona-nos para o tema do risco de crédito conferindo um maior rigor e apuramento do ativo ponderado pelos riscos (RWA) e aos requisitos de capital sobre o risco da contraparte.

*Core Tier 1* / (Risco de Crédito + Risco de Mercado + Risco Operacional)

Nesta nova abordagem inseriram-se novas regras que visam aumentar os requisitos de capital para operações de ‘*trading book*’ e re-securitizadas identificando instrumentos fora do balanço. O objetivo é incluir novas ponderações para a exposição do risco de crédito.

Estas são as alterações mais expressivas que foram efetuadas. Em Basileia III, passará a ser considerado apenas o *Core Tier I* correspondendo ao capital social com as reservas acumuladas. (Ibid.)

### **2.2.1 Novas Medidas – Não constantes em Acordos anteriores**

#### *Capital Conservation Buffer*

Este tipo de capital, visa a obtenção de um reforço (reservas de capital) constituído como auxílio em prováveis momentos de crises financeiras e, desta forma, não haver a necessidade de recorrer a ajudas do Estado.

Segundo BCBS, o *Capital Conservation Buffer* será de 2,5% em 2019, o que fará com que eleve o requisito mínimo de *core Tier I* para 7% no mesmo ano. Para a elaboração do *buffer*, o comité estabeleceu limites à distribuição de lucros, variando de acordo com o nível de capital principal. Com este requisito, o capital total mínimo em 2019 passará a ser de 10,5% em vez dos 8%. Os bancos poderão utilizar este capital para situações que estejam previamente previstas.

#### Counter – Cyclical Capital Buffer

Tem como finalidade a redução da ciclicidade excessiva, isto é, promover um provisionamento prospetivo e conservar capital quando o banco está em crescimento, para na eventualidade de períodos de contração poder ser utilizado, minimizando os riscos inerentes à instituição bancária. Relaciona ainda, o volume de crédito concedido e o PIB de cada país.

Este *buffer* é constituído principalmente por *common equity*. Deve ser composto por uma percentagem adicional que varia entre os 0% e 2,5% do risco ponderado em termos do capital principal. Acaba por ser um capital adicional ao capital de conservação e assim tem-se um capital regulamentado entre 10,5% e os 13% do ativo ponderado pelo risco.

#### Alavancagem

O índice de alavancagem tem que representar 3% do capital principal. O objetivo desta meta é minimizar o endividamento do setor, prevenindo quebras no sistema financeiro. Assim, há uma elevação do capital próprio (património líquido).

#### Liquidez

Através do BCBS, a liquidez passa a ter um papel preponderante para o controlo da solvabilidade nas instituições. Desse modo, introduz dois rácios de ajuda à monitorização do risco de liquidez, sendo:

- *Liquidity Coverage Ratio* (LCR): [ativos de liquidez elevada/fluxos de caixa líquidos nos próximos 30 dias<sup>10</sup>]  $\geq 100\%$
- *Net Stable Funding Ratio* (NSFR): ou rácio de liquidez a longo prazo, pretende monitorizar o prazo dos ativos para garantir que os ativos com longa maturidade tenham

---

<sup>10</sup> Compreendem empréstimos do Estado, dívida soberana sem risco ou com risco reduzido, títulos de dívida privada com desconto; considera-se a diferença entre os fluxos de caixa de saída e de entrada de fundos.

uma potencial fonte estável de *funding*<sup>11</sup> que garanta a liquidez dos bancos. Traduz-se da seguinte forma: [Fundos estáveis disponíveis a longo prazo/Fundos estáveis necessários a longo prazo<sup>12</sup>]  $\geq 100\%$

### Monitorização

Considera-se uma das medidas mais importantes para a estabilidade financeira no setor bancário, sendo que o novo Acordo Basileia III estabeleceu novas medidas que permitem o apuramento do estado financeiro dos bancos em termos de liquidez, e em casos que seja ocorrida alguma situação de alerta, seja visualizada e corrigida de imediato. As novas ferramentas são:

- Mapa de incompatibilidade de fluxos contratuais por maturidades;
- Concentração de financiamento;
- Mapa de ativos disponíveis isentos de encargos;
- Rácio de cobertura de liquidez expresso em moeda relevante;
- Ferramentas de monitorização relacionadas com o mercado.

### Risco de Mercado

O novo acordo não deixou de lado este tipo de risco, visando incluir novas medidas que não ponham em causa o normal funcionamento da instituição, como tal:

- *Buffer* de capital anti cíclico;
- Rácio de endividamento;
- Desenvolvimento de um método que analise o risco sistémico, de cada banco, com base em indicadores quantitativos e qualitativos.

Neste seguimento, foi criado o Conselho Europeu de Risco Sistémico (CERS) com o objetivo de avaliar a estabilidade do sistema financeiro na União Europeia.

Sintetizando, conforme descrito no comunicado do Banco de Portugal (relativo ao aviso do BdP sobre reforço do rácio '*Core Tier 1*' das instituições de crédito), a BCBS visa aumentar a resiliência do setor bancário, através do reforço da qualidade e da consistência do capital

---

<sup>11</sup> Ato de fornecer recursos financeiros, nomeadamente, em forma de dinheiro com a finalidade de financiar uma necessidade.

<sup>12</sup> Inclui capital, ações com maturidade superior a 1 ano, proporção de depósitos esperados que se manterão em caso de *stress test* e passivos com uma maturidade superior a um ano.

regulamentar, com vista a assegurar que os riscos assumidos pelos bancos se encontram adequadamente suportados por uma base de capital de elevada qualidade, que seja facilmente comparável entre instituições. Neste sentido, considerando que a forma predominante de capital deve corresponder a ações ordinárias, reservas e resultados retidos, o BCBS reformulou integralmente as componentes de fundos próprios, que passaram a corresponder ao conjunto de: *Tier 1 capital (going-concern capital)*, ou capital regulamentar com capacidade de absorção de prejuízos numa perspectiva de continuidade da atividade de uma instituição), constituído pelo «*Common Equity Tier 1*» e «*Additional Going-Concern Capital*»; e *Tier 2 capital (gone-concern capital)*, ou capital regulamentar disponível para absorver prejuízos numa perspectiva de liquidação de uma instituição).»

## 2.3 Crédito à Habitação

«O crédito à habitação é um contrato de empréstimo da instituição de crédito ao cliente, por um período de tempo previamente estabelecido, utilizado para aquisição, construção e realização de obras em habitação própria permanente, secundária ou para arrendamento, bem como para aquisição de terrenos para construção de habitação própria.»

(Portal do Cliente Bancário, BdP)

Para contratar um crédito à habitação, ao pretender adquirir um imóvel, o primeiro passo é recorrer a uma instituição de crédito. «Este empréstimo é dos que deixa o banco menos vulnerável, pelo facto de ser garantido pela hipoteca<sup>13</sup>». (Caderno Crédito Habitação, BdP)

Existem três tipos de regimes deste tipo de crédito: - O Geral de Crédito, o Crédito Bonificado e ainda o Crédito Jovem Bonificado. Atualmente, apenas o Regime Geral de Crédito está em vigor.

### 2.3.1 No que consiste a apreciação de um Crédito Hipotecário?

Todo o tipo de crédito antes de ser concedido tem de ser analisado. Esta análise é realizada com outro tipo de cuidado, sendo que também é um tipo de crédito de extrema importância por ser dos mais longos (prazo alargado). De acordo com o tipo de análise efetuada à proposta de crédito habitação, sendo cliente do banco ou não, são solicitados documentos ao cliente para formalizar o crédito. Estes documentos consistem na declaração de rendimentos,

---

<sup>13</sup> Garantia real que confere ao credor (instituição mutuante) o direito de lhe ser pago o montante de garantia, com preferência sobre outros credores. Para além de garantir o capital, garante também a cobrança de juros, juros de mora, e ainda custos decorrentes do eventual processo judicial em caso de incumprimento.

recibos de vencimento, listagem de encargos, entre outros. Se todos estes documentos mostrarem viabilidade presume-se que o cliente estará apto para cumprir com os encargos propostos e procede-se à avaliação da eventualidade do cliente poder entrar em incumprimento.

Os indicadores mais importantes na avaliação de uma operação de crédito hipotecário (CH) são:

- A medição da taxa de esforço, ou seja, a capacidade do cliente em fazer face às suas responsabilidades futuras (encargos/rendimentos). De acordo com a maioria das análises efetuadas, estima-se por prudência que a taxa de esforço não exceda os 30%.
- O valor do rácio *loan to value* (LTV), ou rácio financiamento/garantia (F/G), este indicador evidencia o montante que pode ser financiado, tendo como base o valor de avaliação do imóvel. De acordo com a maioria das análises efetuadas, estima-se por prudência que o LTV não exceda os 80%, sendo este o limite (Frost, 2004 *apud* Barreira, 2012)<sup>14</sup>.

## 2.4 Risco de Crédito

Segundo Neves (2012) o risco de crédito está diretamente ligado à possibilidade de perda, pois quanto maior for essa possibilidade maior será o risco, conseqüentemente.

Para tentar minimizar esse risco de crédito, desenvolvem-se modelos estatísticos na ajuda da tomada de decisão. As probabilidades passam a ter um papel preponderante nas análises estatísticas, sendo que, como menciona Neves (2012:401)

[a]s probabilidades podem ser utilizadas para quantificar o risco de uma aplicação. A média ou valor esperado representa a rendibilidade potencial da aplicação. A variância e desvio-padrão são as medidas estatísticas mais utilizadas na quantificação do risco, sendo frequente usar o coeficiente de variação nas situações em que a rendibilidade média esperada dos ativos que se pretendem comparar sejam diferentes.

Através dos Modelos de Avaliação do Risco (MAR) do Banco de Portugal, (2007) tem-se que o Risco de crédito não é mais do que a probabilidade de ocorrências negativas no capital devido ao incumprimento financeiro de uma contraparte perante a instituição. O risco de crédito existe principalmente nas exposições em crédito (incluindo o titulado), linhas de crédito, garantias e derivados.

---

<sup>14</sup> Frost, S. M. (2004) *The Bank Analyst's Handbook*. West Sussex, England: John Wiley & Sons Ltd.

Geralmente os bancos devem-se concentrar nos três tipos principais de risco: crédito, mercado e operacional. O risco de crédito é a perda potencial que um banco sofreria se não for reembolsado o montante emprestado aos mutuários, de acordo com os termos acordados. O risco de crédito é o maior risco enfrentado pelos bancos e decorre da possibilidade de que os empréstimos ou obrigações não serem reembolsados, parcial ou totalmente (Donko e Herzegovina, 2013).

O risco está presente na maioria das instituições financeiras através das atividades normais de concessão de crédito, pelo facto de inúmeras instituições não terem devidamente definido uma estratégia de análise sólida. Com efeito, todas as decisões creditícias passam pela revisão dos analistas para avaliação dos demais riscos do cliente. Contudo, futuramente não é possível antever ou medir esse risco. Todas as entidades financeiras deverão ter como princípio a concessão de crédito sustentável, visão a longo prazo, com o objetivo de minimizar perdas, e maximizar lucros (Jacobson e Roszbach, 2003).

Os mesmos autores defendem por isso a concessão de grandes empréstimos, de valores avultados de modo a minimizar o risco (através do prazo alargado que se poderá praticar) e na obtenção de fortes ganhos pelos grandes montantes a serem contratados.

### **Qualidade das análises de Crédito**

Há 20 anos, as Instituições Financeiras (IF) baseavam-se, exclusivamente, na análise subjetiva ou nos chamados analistas “*Expert*”, sistemas de avaliação de risco de empréstimos corporativos. Basicamente, esses analistas utilizavam informações sobre os mutuários para avaliação do risco, nomeadamente características como perfil, Capital/endividamento na banca, capacidade/volatilidade dos rendimentos obtidos e garantias, estes eram chamados os pontos fortes para uma análise que seria subjetiva (Altman e Saunders, 1998).

Sommerville e Taffler (1995), defendem que na generalidade os analistas tendem a ser demasiado pessimistas, no qual os sistemas de *Scoring* se tornam um ponto fulcral para mensuração do risco de crédito.

### **Default/Incumprimento**

Incumprimento ocorre quando um devedor não efetua o pagamento programado no fim do 1º período de amortização do empréstimo. Ou seja, um empréstimo está em *default*, assim que o devedor não efetuar o pagamento agendado. Neste contexto, reserva-se o termo “*default*” para qualquer uma das três situações seguintes:

- Um credor foi obrigado a executar uma hipoteca para ganhar o título da propriedade garante do empréstimo;
- O mutuário vende a casa e faz menos do que o pagamento integral na obrigação de hipoteca;
- O credor concorda em negociar ou modificar os termos do empréstimo e perdoa alguns ou todos os principais e juros vencidos. Modificações de empréstimo pode assumir muitas formas, incluindo uma mudança na taxa de juro sobre o empréstimo, uma extensão da duração do empréstimo, e/ou um ajuste do saldo principal devido (Avery, Bostic, Calem e Caner, 1996).

## 2.5 Modelos de *Credit Scoring*

*Scoring* de crédito foi introduzido pela primeira vez, na década de 1940, e ao longo dos anos, evoluiu e desenvolveu-se significativamente. Na década de 1960, com a criação de cartões de crédito, bancos e outros emissores perceberam as vantagens de *scoring* no processo de concessão de crédito.

Na década de 1980, a pontuação de crédito foi utilizado para outros fins, tais como na decisão de aprovação dos pedidos de empréstimo pessoal. Nos últimos anos, a pontuação de crédito tem sido utilizada para os empréstimos à habitação, empréstimos para pequenas empresas e aplicações de seguros e renovações (Koh, Tan, e Goh, 2004; Thomas, 2000 *apud* Yap, Ong e Husen, 2011)<sup>15</sup>.

Os Modelos de *Credit Scoring* são comumente estruturados com base nos moldes do modelo de 1968, de *Z-score de Altman*. Os *Credit Scoring Models* (CSM) utilizam dados históricos, combinados com uma técnica estatística - um modelo linear de probabilidade, um modelo de regressão *logit* ou *probit* ou análise discriminante múltipla, a fim de identificar quais as características eficazes para distinguir entre mutuários cumpridores e não cumpridores de crédito. Numa segunda fase, o modelo é usado para calcular uma pontuação para cada novo candidato de empréstimo. Prevê-se que uma maior pontuação indique um melhor desempenho esperado do mutuário e, portanto, uma menor probabilidade de *default* (PD). Esta pontuação deve ser comparada a um ponto de corte para determinar se o pedido for aceite/rejeitado

---

<sup>15</sup> Koh, H. C., Tan, W. C., & Goh, C. P. (2004). Credit scoring using data mining techniques. *Singapore Management Review*, 26(2), 25–47

Thomas, L. C. (2000). A survey of credit and behavioral scoring: Forecasting financial risk of lending to customers. *International Journal of Forecasting*, 16, 149–172.

ou, por outra, requer uma avaliação mais aprofundada. Assim, uma calibração do modelo é necessária de modo a que seja determinado o ponto ótimo de corte (Thanh e Kleimeier, 2006).

Também segundo Mester (1997), para construir um modelo de pontuação, ou “*scorecard*”, deverão ser analisados os dados históricos sobre o desempenho dos empréstimos feitos anteriormente, para determinar quais as características do mutuário, que serão úteis para revelar se o empréstimo tem um bom desempenho. Um modelo bem projetado deve traduzir uma maior percentagem de pontuações elevadas para os mutuários, cujos empréstimos tenham um bom desempenho, e, menor percentagem para os mutuários cujos empréstimos não terão tão bom perfil.

Quanto maior a pontuação de crédito, maior a qualidade de crédito de um candidato, enquanto o menor a pontuação de crédito, menor a qualidade de crédito de um candidato (Samreen e Zaidi, 2012).

Para avaliar o risco de crédito, os credores devem reunir informações sobre uma série de fatores, incluindo o presente e o passado da situação financeira do mutuário e da natureza e valor do imóvel que serve como garantia do empréstimo (Avery, Bostic, Calem, Canner e Matson, 1996).

Um modelo de *credit scoring* fornece uma estimativa do risco de crédito do mutuário - ou seja, a probabilidade do mutuário pagar o empréstimo, como contratado, com base em uma série de características quantificáveis respeitantes ao mesmo (Dinh e Kleimeier, 2007).

O *Scoring* assume que o desempenho de empréstimos futuros com um conjunto de determinadas características será o reflexo do conjunto de empréstimos passados com características semelhantes. Um modelo de *credit scoring* é uma fórmula que coloca pesos nas diferentes características de um mutuário, um credor e de um empréstimo. A fórmula produz uma estimativa da probabilidade ou, ainda, o risco de que determinado resultado poderá acontecer (Schreiner, 2000).

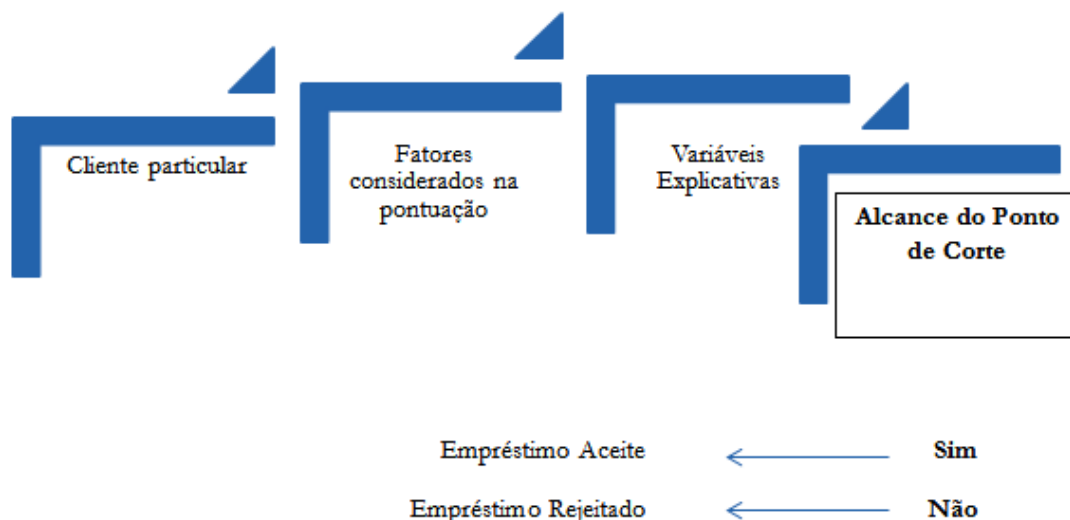
Estes modelos de classificação baseados na pontuação de crédito analisam as características dos indivíduos, como idade, renda, estado civil, histórico de pagamento, entre outros, são utilizados para classificar os novos clientes em bons ou maus pagadores (Chen e Huang, 2003 *apud* Samreen e Zaidi, 2012)<sup>16</sup>.

---

<sup>16</sup> Chen, M., & Huang, S. (2003). Credit Scoring and Rejected Instances Reassigning Through Evolutionary Computation Techniques. *Expert Systems with Applications*, 24, 433-441.

Ainda segundo os mesmos autores muitos bancos utilizam apenas dois grupos de classificação como “bom” ou “mau” candidatos; e muitos usam três grupos “bom”, “mau” ou “recusado”. Cabe aos gestores de crédito analisarem os pedidos indeferidos.

O processo do *Credit Scoring*, sintetizado, será o seguinte:



### 2.5.1 Seleção do Modelo

São pelo menos conhecidas 4 abordagens metodológicas para o desenvolvimento de um sistema multivariado de pontuação de crédito: i) modelo de probabilidade linear; ii) o modelo *logit*; iii) o modelo *probit*; e, por fim, iv) modelo de análise discriminante.

Martin (1997) utilizou nas suas pesquisas o modelo *logit* e a análise discriminante, para prever o *Fail bank* em 1975, período em que 23 bancos faliram. Ambos os sistemas transmitiram semelhantes conclusões em termos de identificação de fracassos/lacunas. West (1985)<sup>17</sup> utilizou o modelo *logit* com o auxílio da análise fatorial para medir a condição de probabilidade de *Fail* de um banco. O que com a investigação efetuada, os fatores identificados pelo modelo *logit*, curiosamente são os mesmos utilizados pelos examinadores bancários. Platt e Platt (1991)<sup>18</sup> utilizaram o modelo *logit* para testes de índices contabilísticos para a indústria e denotou-se uma melhor identificação dos fatores de falência também nesse setor. No mesmo

<sup>17</sup> West, R.C., 1985. A factor-analytic approach to bank condition. *Journal of Banking and Finance* 253-266.

<sup>18</sup> Platt, H.D., Platt, M.B., 1991b. A linear programming approach to bond portfolio selection. *Economic Financial Computing*, Spring 71-84.

seguimento, Lawrence et al. (1992)<sup>19</sup> concluíram que o histórico do cumprimento dos créditos está longe de ser o ponto-chave na previsão de incumprimentos futuros.

A inclusão de novas variáveis macroeconómicas causam algum impacto no sistema aplicativo podendo assim conduzir a novos modelos de crédito. Da mesma forma que a envolvente está sujeita a alterações económicas, as variáveis utilizadas podem seguir a mesma linha comprometendo assim o comportamento das mesmas consoante essas alterações/mudanças (Maddala, 1983 *apud* Greene, 1998)<sup>20</sup>.

Os ajustes efetuados no modelo são com base numa razão de verossimilhança, log-likelihood ratio (LLR) derivada a partir do procedimento de máxima probabilidade utilizada para definir o modelo. Para cada variável macroeconómica incluída, a relação que dá o valor de  $p$  menor está incluído num modelo ideal. Há que ter presente que, tendo um elevado número de testes e de variáveis para testar, o tempo de processamento para descodificar cada modelo é longo, com isto consegue-se excluir um elevado número de processos pelo demasiado tempo que levam na demonstração de resultados (Hosmer e Lemeshow, 1999 *apud* Bellotti e Crook, 2007)<sup>21</sup>.

Três exigências principais são considerados para a validação de um modelo de pontuação de crédito: estabilidade, legibilidade e poder discriminatório.

Estabilidade. Um modelo estável requer coeficientes bem determinados, com alta confiança e resultados semelhantes sobre as características de desempenho se testado dentro e fora da amostra.

Legibilidade. Um modelo é dito legível quando os seus coeficientes podem ser facilmente interpretado.

Poder discriminatório. Este é definido pelo CBSB (2005) como a capacidade de classificar corretamente observações sobre a base de probabilidade de incumprimento, atribuindo pontuações (Van Gestel et al., 2006 *apud* Gool et al., 2009)<sup>22</sup>.

---

<sup>19</sup> Lawrence, E.L., Smith, S., Rhoades, M., 1992. An analysis of default risk in mobile home credit. *Journal of Banking and Finance* 299±312.

<sup>20</sup> Maddala, G., 1983. *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York.

<sup>21</sup> Hosmer Jr. DW and Lemeshow S (1999). *Applied Survival Analysis: regression modelling of time to event data*. Wiley.

<sup>22</sup> Van Gestel, T., Baesens, B., Van Dijke, P., Garcia, J., Suykens, J. and Vanthienen, J. (2006). *A process model to develop an internal rating system: Sovereign credit ratings*. *Decision Support Systems*, 42(2), pp 1131-1151.

## 2.6 Credit Scoring

O *Credit Scoring* tem dado uma enorme relevância aos processos de decisão de crédito, devendo-se o sucesso aos aplicativos, alterando totalmente o *BackOffice* das instituições. O *Credit Scoring* não é mais do que uma análise à qualidade do crédito, ou seja, consiste em medir o risco de crédito de um determinado número de créditos concedidos através dos dados do perfil dos contraentes. Esses dados desempenham o papel de variáveis do empréstimo, permitindo a construção do modelo que calcule a probabilidade de incumprimento mediante essas variáveis creditícias (Raymond, 2007 *apud* Semedo, 2009)<sup>23</sup>.

Segundo Neves (2012) os sistemas de análise de risco de crédito podem classificar-se em dois tipos, os de *scoring* (pontuação) e modelos de *rating* (notação de crédito). Sendo que os sistemas de *scoring* aplicam indicadores que permitem classificar o cliente de forma automatizada, gerando uma pontuação através de informação histórica.

Ainda segundo o mesmo autor, os modelos de *scoring* podem ser de natureza teórica ou empírica, sendo que os primeiros, respetivamente, têm como base de análise a experiência e sensibilidade do analista de crédito bem como dos seus conhecimentos, embora esta metodologia esteja fora do âmbito do que é exigido segundo a Supervisão e Acordos de Basileia II e III. Para o mesmo autor (2012:444):

Estes métodos têm a vantagem da fácil e rápida concessão, mas, porque a realidade é complexa, a percepção teórica e a sensibilidade do analista financeiro podem conduzir a um modelo ineficaz na seleção e classificação das empresas no seu grau de risco. Pior ainda quando, na prática, raramente existe controlo *a posteriori* da eficiência do sistema. Este sim, deveria ser uma exigência da Supervisão, de testar a qualidade dos sistemas de análise de risco de crédito como um *benchmarking* nacional.

Através de práticas estatísticas, os modelos empíricos conseguem-nos dar os rácios mais adequados e as ponderações mais acertadas, de modo a classificar um cliente com um grau de risco fiável. As práticas estatísticas dizem-se paramétricas ou não paramétricas. Das primeiras, respetivamente, tem-se as análises univariadas e multivariadas, onde nas multivariadas a análise discriminante (*logit e probit*) é a mais utilizada na análise de risco de crédito (Ibid.).

---

<sup>23</sup> Raymond, Anderson. *The Credit Scoring Toolkit Theory and Practice for Retail Credit Risk Management and Decision Automation*. New York: OXFORD University Press Inc., New York, 2007.

«A credit scoring model provides an estimate of a borrower's credit risk – i.e. the likelihood that the borrower will repay the loan as promised, based on a number of quantifiable borrower characteristics.» (Dinh e Kleimeier, 2007)

*Credit scoring is based on statistical or operational research methods. Historically, discriminant analysis and linear regression have been the most widely used techniques for building scorecards. Other techniques include logistic regression, probit analysis, nonparametric smoothing methods especially j-nearest neighbors, mathematical programming, Markov chain models, recursive partitioning, expert systems, genetic algorithms, Markov chain models, recursive partitioning, expert systems, genetic algorithms and neural networks (Hand e Henley, 1997 apud Yap, Ong e Husain, 2011: 13275)<sup>24</sup>.*

O paradigma em que se baseia o *credit scoring* é na aprendizagem dos clientes atuais, na identificação das características dos bons clientes, e na concessão de crédito a novos clientes com características semelhantes às daquelas. O método produz um “*score*” que o banco utiliza para posicionar o crédito em termos de risco e decidir quanto à concessão, ou não do crédito (Loretta, 1997 apud Pereira e Chorão, 2007)<sup>25</sup>.

*Credit Scoring* é o conjunto de modelos de decisão e de técnicas implícitas que os credores se baseiam como ferramenta na concessão de crédito ao consumidor.

Estas técnicas decidem quem está apto a aceder ao crédito, qual o valor contratado, e que estratégias operacionais irão melhorar a rentabilidade do investimento efetuado pelo mutuário para o credor (Thomas, Edelman e Crook, 2002).

É um método de avaliação do risco de crédito, através de uma pontuação atribuída aos empréstimos requeridos com base nas diferentes características de um mutuário, credor e empréstimo. A fórmula produz uma estimativa da probabilidade ou do risco de que um determinado resultado ocorrerá. A Avaliação de crédito é um dos processos mais importantes nas instituições bancárias e, conseqüentemente nas “decisões de gestão de crédito” (Chakravarty e Jha, 2012).

Conforme Mester (1997) e Hansell (1995), usando dados históricos e técnicas estatísticas, o *credit scoring* tenta isolar os efeitos de várias características recorrentes sobre *default*/incumprimento. O método produz um *score*, o qual o banco pode usar para classificar os proponentes do empréstimo ou mutuários em termos de risco.

---

<sup>24</sup> Hand, D. J., & Henley, W. E., (1997). Statistical classification methods in consumer credit scoring: A review Journal of the Royal Statistical Society: Series A. Statistics in Society, 160 (3), 523-541.

<sup>25</sup> LORETTA, J. M. (1997), “What’s the point of Credit Scoring”, *Business Review*, Federal Reserve Bank of Philadelphia.

Este conceito foi utilizado em 1950, inicialmente direcionado para o segmento dos cartões de crédito, atualmente cada vez mais é utilizado no crédito hipotecário (Mester, 1997).

Este processo inclui a recolha, análise e classificação dos diferentes elementos de crédito e variáveis para avaliar decisões. A qualidade dos empréstimos bancários é o principal determinante da concorrência, da sobrevivência e lucratividade dos mesmos. Um dos 'kits' mais importantes para classificar os clientes como parte do processo de avaliação de crédito para reduzir o atual risco esperado de um 'mau' cliente, é a pontuação de crédito (Hand e Jacka, 1998 *apud* Abdou e Pointon, 2011)<sup>26</sup>.

Consequentemente, a pontuação de crédito, pode ser simplesmente definida como o uso de modelos estatísticos, para transformação de dados relevantes em medidas numéricas que norteiam as decisões de crédito. É a industrialização de confiança; um futuro desenvolvimento lógico das notações de risco (Anderson, 2007 *apud* Abdou e Pointon, 2011)<sup>27</sup>.

Thomas (2000) explica as diferenças entre dois tipos de modelos de crédito, sendo os de âmbito comportamental e os de aprovação de crédito propriamente ditos. No entanto, há um outro modelo, nomeadamente, com base num histórico ou *Behavioural Scoring*, que explicam as relações de créditos já existentes, ou seja, para clientes que já têm relacionamento creditício com o banco. A finalidade dos *Behavioural Scoring* é estimar a probabilidade do cliente poder ficar em incumprimento. Os modelos de aprovação de crédito são ferramentas que dão suporte à tomada de decisão na concessão de crédito, com o objetivo em estimar a probabilidade de incumprimento de um novo mutuário.

No setor financeiro, os consumidores induzem-se regularmente ao crédito para satisfazerem as suas necessidades. O risco para as instituições financeiras, na concessão do crédito solicitado depende de quão bem eles consigam distinguir os bons dos maus clientes de crédito. Uma técnica amplamente adotada para resolver este problema designa-se então, "*Credit Scoring*".

*Credit Scoring* não é mais do que um procedimento informatizado que prevê a atribuição de uma pontuação a cada mutuário com base em dados característicos do seu perfil como rendimento, profissão, idade, etc.,. Caso a pontuação seja superior ao ponto de corte fixado pelo credor estão reunidas as condições para concessão do crédito, caso contrário o empréstimo será rejeitado (Steenackers e Goovaerts, 1989 *apud* Nguyen, 2014)<sup>28</sup>.

---

<sup>26</sup> Hand, D. J., Jacka, S. D. 1998. *Statistics in Finance*, Arnold Applications of Statistics: London.

<sup>27</sup> Anderson, R. 2007. *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. New York: Oxford University Press.

<sup>28</sup> Steenackers A. and Goovaerts M.J. (1989), *A Credit Scoring Model for Personal Loans*, *Insurance: Mathematics and Economics* 8 (8): 31-34.

Além disso, ele ajuda a avaliar o risco do empréstimo. O método de *Scoring* é uma avaliação confiável da qualidade de crédito de uma pessoa, uma vez que se baseia em dados reais.

Segundo Muller e Hardle (2002) os métodos de pontuação de crédito destinam-se a estimar o risco de incumprimento de possíveis mutuários e classificá-los em grupos de acordo com seu nível de risco. Isto envolve normalmente o cálculo de uma pontuação de crédito e a estimativa da PD.

Ainda de um ponto de vista estatístico, conforme os mesmos autores, a classificação entre detentores de risco e ‘não risco’ é acima de tudo um problema de análise discriminante. Soluções clássicas para isso são a análise discriminante linear ou quadrática e - em um nível mais avançado – Análise discriminante e Análise Logística. Todos estes métodos têm como base uma pontuação dependendo das variáveis explicativas.

### Objetivos do *Scoring*

O objetivo principal dos métodos de *scoring* está em melhorar o processo de identificação e seleção dos bons clientes, com intuito de reduzir o incumprimento e, por conseguinte, redução de perdas. Torna-se assim fácil entender que o *scoring* é uma ferramenta imprescindível nas decisões inerentes à gestão do risco.

Segundo Semedo (2009) os *scores* gerados pelo modelo calculam a perda máxima (*expected loss*) da carteira de crédito da instituição e, conseqüentemente determina a provisão necessária de forma a cobrir a perda máxima esperada.

Pelo mesmo autor (2009:20)

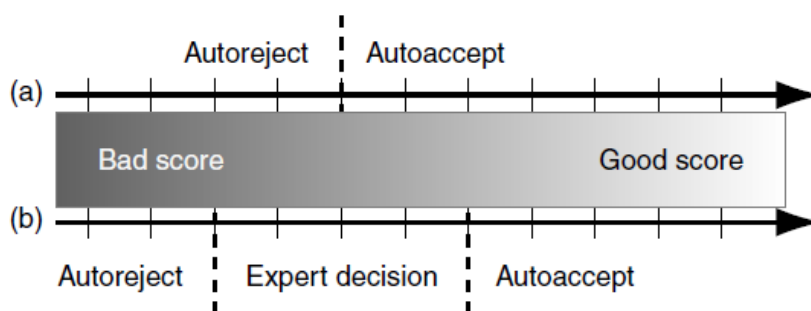
[p]ara determinar a perda esperada, o risco de incumprimento da carteira de crédito precisa ser adequadamente quantificado e os *scores* têm demonstrado ser um importante *input*. Os scores são igualmente utilizados para determinar o montante de capital necessário para proteger as instituições financeiras e os depositantes de perdas inesperadas – capital económico/regulamentar.

Outras das áreas onde o *scoring* pode ser igualmente utilizado é no decisor de *pricing*: *risk-based pricing* (ou *risk-adjusted pricing*). Neste caso, a finalidade é determinar o preço do produto requerido pelo cliente através do perfil de risco do mesmo, de acordo com o *score* que lhe é atribuído.

### Aplicação de um sistema de *Score*

Sistemas de pontuação têm como base dois cenários. Em sistemas de pontuação de aplicação totalmente automatizados, o pedido é aprovado se a pontuação for superior ao valor limite ou ponto de corte; e, em sistemas semiautomáticos, o pedido é automaticamente aprovado ou rejeitado por pontuações altas e baixas. Para pontuação intermédia a decisão decorre de uma análise adicional pelo analista de crédito.

**Figura 2.7 - Sistema de *Score* de Crédito**



Fonte: Gestel & Baesens, (2009:98)

### 2.6.1 Vantagens e Desvantagens do *Scoring*

Podemos apontar as seguintes vantagens da utilização de modelos de *scoring*:

- Redução do tempo de análise das propostas de crédito;
- Após a automatização do processo, o cálculo dos *scores* torna-se mais rápido, gerando uma decisão de concessão ou rejeição da proposta de crédito em tempo real, sendo fulcral no mercado atual cada vez mais competitivo;
- O *score* torna o processo de decisão objetivo eliminando dúvidas;
- Aumento do lucro da instituição através de uma maior assertividade de propostas aprovada, com reduzidos níveis de *default*;
- Permite o tratamento personalizado do cliente;
- As estratégias de risco/crédito podem ser mais facilmente identificadas pela organização;
- Aumento da qualidade dos serviços prestados aos clientes;
- O processo torna-se de fácil entendimento pelos intervenientes da proposta de crédito;
- Permite discriminar as variáveis inerentes ao incumprimento.

Sintetizando, o *scoring* de crédito, tem de facto algumas vantagens que levaram ao crescimento da utilização deste método na avaliação do empréstimo.

Em primeiro lugar, e de extrema importância, reduz o tempo necessário no processo de aprovação do empréstimo. Um estudo efetuado pelo *Business Banking Board* descobriu que as médias do processo de aprovação de empréstimos tradicionais vão desde 1/2 horas por empréstimo de pequenas empresas, e no passado, o panorama torna-se mais grave com 2 semanas para analisar um empréstimo (Allen, 1995 *apud* Mester, 1997)<sup>29</sup>.

A pontuação de crédito pode reduzir o tempo em menos de uma hora, embora a economia de tempo varie, dependendo se o banco segue rigorosamente o ponto de corte de crédito ou se se reavaliam aplicações com pontuações próximas do ponto de corte. *Barnett Bank*, efetuou um estudo que relata uma diminuição do tempo de processamento em três ou quatro semanas com a utilização de um ponto de corte (Lawson, 1995 *apud* Mester, 1997)<sup>30</sup>.

Há benefícios consideráveis de tempo, se deixarmos o raciocínio humano e darmos possibilidade de novas tecnologias/aplicações tomarem decisões. Haverá uma forma mais rápida de contar os empréstimos “ruins”, minimizando o tempo despendido em cada análise de crédito, fazendo-se assim um maior número de pequenos negócios. O uso da intuição do analista e da experiência refletida poderá fazer com que créditos não sejam efetuados, o que com um modelo de *Credit Scoring* atualizado, isso não acontecerá (Hansell, 1995 New York times). Outro benefício do *scoring* de crédito é que o mesmo requer de menos informação para tomar uma decisão. Os modelos de *scoring* de crédito foram estimados para incluir apenas as variáveis estatisticamente significativas (Crook,1996 *apud* Abdou e Pointon, 2011)<sup>31</sup>.

Denota-se, que a economia de tempo é das maiores vantagens. O apoio para o processo de tomada de decisão torna-se mais rigoroso; há minimização dos custos e esforços de processamento de créditos; menor número de erros cometidos; fornecimento de estimativas para ser comparado em auditorias; inclusão de variáveis suportadas através de análise objetiva para avaliar o risco de crédito; modelos elaborados com base em dados reais; são consideradas as inter-relações entre as variáveis.

Da mesma forma que os autores anteriores mencionam alguns benefícios do uso do *scoring*, também Itoo e Selvarasu (2015) consideram igualmente que o tempo é uma das grandes

---

<sup>29</sup> Allen, James C. “A Promise of Approvals in Minutes, Not Hours,” *American Banker* (February 28, 1995), p. 23

<sup>30</sup> Lawson, James C. “Knowing the Score,” *US Banker* (September 1995), pp. 61-65.

<sup>31</sup> Crook, J. N. 1996. Credit scoring: An overview. Working paper series No. 96/13, British Association, Festival of Science. University of Birmingham, the University of Edinburgh.

vantagens do uso desta técnica, tal como o uso de fatores que são estatisticamente significativos e correlacionados. O mesmo defende ainda que há uma minimização dos custos proveniente da redução do tempo.

Após distinguidas as vantagens, existem contudo alguns inconvenientes associados ao modelo, sendo:

- Custos associados ao desenvolvimento do *software/hardware*;
- Qualidade dos dados das instituições por vezes escassa ou inapropriada à implementação do sistema;
- A implementação deste tipo de modelo altera, em parte, os processos operacionais da instituição;
- Este modelo baseia-se em factos históricos, o que nem sempre é um processo de sucesso;
- É um sistema complexo de possíveis erros no seu desenvolvimento, estratégia ou implementação, e com um eventual erro poderá acarretar danos na instituição de ordem elevada;
- Normalmente, a base de dados de clientes em incumprimento utilizada para testar um modelo desta ordem parte do pressuposto que o número de casos de *default* é suficiente, no entanto, poderá não ser, sendo que cada caso de risco é um caso isolado, por motivos distintos e perfis de clientes por vezes bastante diferentes (Semedo, 2009 *apud* Raymond, 2007)<sup>32</sup>.

A precisão dos sistemas de pontuação ainda é uma questão em aberto. Precisão é uma consideração muito importante na utilização de pontuação de crédito porque mesmo que o credor possa reduzir os seus custos de avaliação dos pedidos de empréstimo usando o *scoring*, se os modelos não forem precisos, estas poupanças seriam devoradas por empréstimos de fraca rendibilidade (Mester, 1997).

A pontuação de crédito pode reduzir custos e trazer uma maior consistência ao processo de subscrição, contudo a sua confiabilidade depende da exatidão, integridade, e tempestividade das informações utilizadas para gerar as pontuações. Por exemplo, se o relatório de crédito apresentar erros ou estiver incompleto não será possível medir com precisão o risco representado por um mutuário, podendo levar a decisões indesejadas numa operação. Ainda há outros fatores que o modelo não consegue abranger, como rendas que o mutuário possa vir a ter futuramente ou não documentadas, entre outros encargos semelhantes de carácter variável.

---

<sup>32</sup> Raymond, Anderson. *The Credit Scoring Toolkit Theory and Practice for Retail Credit Risk Management and Decision Automation*. New York: OXFORD University Press Inc., New York, 2007.

Um modelo de *Scoring* não deve ser estático, deverá ser submetido a testes constantes, de modo a garantir a sua precisão atualizada ao longo do tempo (Avery, et al., 1996).

A pontuação de crédito pode usar qualquer característica de um cliente. Fatores económicos, por vezes, não estão incluídos nas análises. Através de modelos de *scoring* de crédito, os clientes podem ser distinguidos com determinadas características, que os tornam semelhantes aos maus pagadores (em vez dos bons), podendo se tratar de um acaso e acarretar problemas de classificação incorreta. Estatisticamente, um modelo de pontuação de crédito é “incompleto”, pois deixa de fora algumas variáveis que, relacionadas com outras, pode prever o perfil de um bom cliente. Contudo, a menos que um modelo de pontuação de crédito tenha todas as variáveis possíveis incluídas no mesmo, normalmente classificará de forma errada alguns mutuários.

Além disso, os modelos de *scoring* de crédito não são padronizados e diferem de um mercado para outro; são dispendiosos para comprar e, posteriormente, para formar os analistas de crédito; e, por vezes, um sistema de pontuação de crédito pode rejeitar um cliente digno de crédito, por alteração de informação de alguma variável, por exemplo, o local de residência, profissão, etc. (Ibid.).

## **2.7 Algumas Técnicas Estatísticas utilizadas em *Credit Scoring***

A estatística é uma parte da matemática aplicada que fornece métodos para recolha, organização, descrição, análise e interpretação de dados, e a utilização dos mesmos na tomada de decisões. Segundo Reis, Melo, Andrade e Calapez (1996:18) «[p]ara fazer face à mudança, é necessário que as decisões e o planeamento se apoiem numa análise cuidada da situação presente e numa previsão realista do que acontecerá no futuro».

O estudo da estatística, incide em fatores aleatórios e a sua função primordial é a representação e explicação das observações quantitativas numéricas, relativas à resolução de um determinado problema, conforme referido por Murteira, Ribeiro, Andrade e Silva e Pimenta (2010). São casos que envolvem uma multiplicidade de causas, e por fim, são representados sob forma analítica, geralmente através de gráficos, tabelas e quadros estatísticos.

Murteira et al., (2010) referem ainda que o método estatístico tem como base a aplicação da teoria estatística das probabilidades, e a sua aplicação para descrever vários padrões sociais, económicos, entre outros. Reis et al., (1998) reiteram que a sua interpretação deve considerar um determinado grau de precisão, pois não são absolutamente verdadeiras, ainda assim, são bastante aceites por parte dos investigadores com preocupações de ordem quantitativa.

### 2.7.1 Métodos utilizados em *Credit Scoring*

São vários os principais métodos paramétricos e não paramétricos usados em *Credit Scoring*:

- Regressão Linear: a principal técnica são os mínimos quadrados ordinários, adequado em situações em que a variável resposta é contínua;
- Análise Discriminante: a principal técnica é a distância de Mahalanobis, classifica os objetos em grupos pré-definidos, minimizando a variância;
- Regressão Logística: a principal técnica utilizada são estimadores de máxima verossimilhança, adequada em situações em que a variável resposta é binária;
- Árvores de Decisão: a principal técnica é *Chaid*, utiliza a estrutura da árvore para maximizar a variância entre grupos;
- Redes Neurais: a principal técnica é a percepção multicamada, técnica de inteligência artificial, os resultados são difíceis de explicar;
- Programação Linear: a principal técnica é o método simplex, muito utilizado na otimização de alocação de recursos (Raymond, 2007 *apud* Semedo, 2009)<sup>33</sup>.

Como defende Hand e Henley (1997) de modo geral não existe um melhor método. A melhor técnica dependerá dos detalhes do empréstimo, ou seja, a estrutura dos dados, as características usadas e o conjunto de variáveis a ter em conta.

### 2.7.2 Modelos de Regressão

Para tentar minimizar o risco de crédito, desenvolvem-se modelos estatísticos de ajuda às tomadas de decisão. As probabilidades passam a ter um papel preponderante nas análises estatísticas, como menciona Neves (2012:401)

«[a]s probabilidades podem ser utilizadas para quantificar o risco de uma aplicação. A média ou valor esperado representa a rentabilidade potencial da aplicação. A variância e desvio-padrão são as medidas estatísticas mais utilizadas na quantificação do risco, sendo frequente usar o coeficiente de variação nas situações em que a rentabilidade média esperada dos ativos que se pretendem comparar sejam diferentes.»

Através de práticas estatísticas, os modelos empíricos conseguem revelar os rácios mais adequados e as ponderações mais acertadas, de forma a classificar um cliente com um grau de risco de maior fiabilidade. As práticas estatísticas dizem-se paramétricas ou não paramétricas.

---

<sup>33</sup> Raymond, Anderson. *The Credit Scoring Toolkit Theory and Practice for Retail Credit Risk Management and Decision Automation*. New York: OXFORD University Press Inc., New York, 2007.

Das primeiras, respetivamente, tem-se as análises univariadas e multivariadas, onde nas multivariadas a análise discriminante (*logit e probit*) é a mais utilizada na análise de risco de crédito (Ibid.).

Para a criação do modelo em estudo, usamos um modelo estatístico de regressão, a mesma segundo Pestana e Gageiro (2005) tem como objetivo explicar e prever o comportamento da variável Y (variável endógena de natureza quantitativa) em função de uma X (variável exógena de natureza quantitativa), necessitando para tal da expressão analítica que a traduz, obtida através do Método dos Mínimos Quadrados (MMQ).

Existem algumas vantagens do uso do método de regressão logística e não linear, como defende Rosenfeld (2002) torna-se mais fácil adicionar variáveis além das pré-definidas.

### 2.7.3 Regressão Linear e Logística

Como refere Pestana e Gageiro (2005), as associações lineares entre uma variável dicotómica ou categórica (Y) e variáveis explicativas (X), definem-se por modelo estatístico de regressão linear.

A regressão logística descreve-se como uma técnica estatística utilizada com a finalidade de obter um modelo que permita a predição de valores tomados por uma variável categórica, com base num conjunto de observações. A variável categórica tem normalmente uma tipicidade binária a partir de uma série de variáveis explicativas contínuas ou igualmente binárias (Christensen, 1997).

A análise *Logit* ou regressão logística, é, pois, uma técnica estatística utilizada na separação de dois grupos de variáveis com a finalidade de obter a probabilidade de uma observação, pertencente a um determinado conjunto, em função do comportamento das variáveis independentes.

Para a realização de um método como o proposto, a técnica mais utilizada é o método de regressão Logística. Para entender o modelo mais simples, (Wooldridge: 22) refere:

*“Much of applied econometric analysis begins with the following premise:  $y$  and  $x$  are two variables, representing some population, and we are interested in “explaining  $y$  in terms of  $x$ ,” or in “studying how  $y$  varies with changes in  $x$ .” (...) In writing down a model that will “explain  $y$  in terms of  $x$ ,” we must confront three issues. First, since there is never an exact relationship between two variables, how do we allow for other factors to affect  $y$ ? Second, what is the functional relationship between  $y$  and  $x$ ? And third, how can we be sure we are capturing a *ceteris paribus* relationship between  $y$  and  $x$  (if that is a desired goal)? We can resolve these ambiguities by writing down an equation relating  $y$  to  $x$ . A simple equation is:*

$$y = \beta_0 + \beta_1 x + u. \quad (2.1)$$

Equation which is assumed to hold in the population of interest, defines the simple linear regression model. It is also called the two-variable linear regression model or bivariate linear regression model because it relates the two variables  $x$  and  $y$ .

When related by equation, the variables  $y$  and  $x$  have several different names used interchangeably, as follows.  $y$  is called the dependent variable, the explained variable, the response variable, the predicted variable, or the regressand.  $X$  is called the independent variable, the explanatory variable, the control variable, the predictor variable, or the regressor. (The term covariate is also used for  $x$ .) The terms “dependent variable” and “independent variable” are frequently used in econometrics. But be aware that the label “independent” here does not refer to the statistical notion of independence between random variables.

The terms “explained” and “explanatory” variables are probably the most descriptive. “Response” and “control” are used mostly in the experimental sciences, where the variable  $x$  is under the experimenter’s control. We will not use the terms “predicted variable” and “predictor,” although you sometimes see these. Our terminology for simple regression is summarized in Table”

**Figura 2.8 - Terminology for Simple Regression**

$y$	$x$
Dependent Variable	Independent Variable
Explained Variable	Explanatory Variable
Response Variable	Control Variable
Predicted Variable	Predictor Variable
Regressand	Regressor

**Fonte:** Adaptado de Wooldridge (2002:23)

«The variable  $u$ , called the error term or disturbance in the relationship, represents factors other than  $x$  that affect  $y$ . A simple regression analysis effectively treats all factors affecting  $y$  other than  $x$  as being unobserved. You can usefully think of  $u$  as standing for “unobserved.”»

Deste modo, através do modelo de regressão simples, consegue-se explicar como usar a variável dependente  $y$  em função de uma variável independente  $x$ . Seguindo esta forma, encon-

tra-se alguma desvantagem, porque para trabalhos empíricos, não se retiram as devidas conclusões, acerca da relação da variável dependente  $y$  com a independente  $x$ , *ceteris paribus*<sup>34</sup>. Uma vez que o trabalho a ser desenvolvido está baseado num estudo empírico, o método que dará as melhores conclusões será a Regressão Múltipla. Assim, (Wooldridge: 66) defende que:

*“Multiple regression analysis is more amenable to ceteris paribus analysis because it allows us to explicitly control for many other factors which simultaneously affect the dependent variable. This is important both for testing economic theories and for evaluating policy effects when we must rely on nonexperimental data. Because multiple regression models can accommodate many explanatory variables that may be correlated, we can hope to infer causality in cases where simple regression analysis would be misleading.*

*Naturally, if we add more factors to our model that are useful for explaining  $y$ , then more of the variation in  $y$  can be explained. Thus, multiple regression analysis can be used to build better models for predicting the dependent variable.*

*An additional advantage of multiple regression analysis is that it can incorporate fairly general functional form relationships. In the simple regression model, only one function of a single explanatory variable can appear in the equation. As we will see, the multiple regression model allows for much more flexibility.”*

A regressão logística é uma técnica paramétrica padrão utilizada para a pontuação de crédito, ou seja, para a obtenção de um *score*. Ele é utilizado para modelar o *log*-probabilidades de um evento dado um vetor de  $n$  variáveis  $x$ :

$$\log\left(\frac{p}{1-p}\right) = \mathbf{w} \cdot \mathbf{x} \quad (2.2)$$

Em que  $p$  é a probabilidade do evento e  $w$  é uma combinação linear das ponderações em  $x$ . Tendo um conjunto de  $n$  testes de observações  $\mathbf{x}_1$  para  $\mathbf{x}_n$ , o vetor de maior ponderação pode ser estimado usando estimativa de probabilidade máxima. Para a pontuação de crédito, o evento é o padrão, portanto,  $p$  é o PD (Thomas, et al., 2002 *apud* Bellotti e Crook, 2008)<sup>35</sup>.

<sup>34</sup> Expressão do latim que significa “mantendo nas mesmas condições”, ou seja, explica a influência de um fator sobre outro, sem as variáveis em uso sofrerem alterações.

<sup>35</sup> Thomas L. C., Edelman DB and Crook JN (2002). *Credit Scoring and its Applications*. SIAM Monographs on Mathematical Modeling and Computation. SIAM: Philadelphia, USA.

Than e Kleineier (2006) defendem, igualmente, que o melhor modelo a ser utilizado é o de regressão logística. Com base em uma amostra de empréstimos existentes (histórico), as características do mutuário serão, posteriormente, observadas e tidas em conta, bem como outros fatores.

O problema passa pela escolha de variáveis não observáveis, por exemplo, um mutuário que tenha tido crédito vencido na banca, ao visualizarmos esse facto não podemos medir diretamente a PD do mutuário através dessa informação.

A técnica de regressão logística supera esse problema, estimando diretamente essa probabilidade e, portanto, tem sido a metodologia de escolha para uma grande maioria em técnicas de risco de crédito (Schreiner, 1999 *apud* Than e Kleineier, 2006)<sup>36</sup>.

Esta técnica assume a existência de uma variável contínua  $Z_j$  que é definida como a probabilidade de que os padrões candidatos  $j$  possam ser modelados como uma função linear de um conjunto de variáveis  $x$  (Than e Kleineier, 2006):

$$Z_j = W'x = w_1x_{j1} + w_2x_{j2} + \dots + w_kx_{jk} \quad (2.3)$$

A técnica mais comum usada para a classificação de crédito relacionado com a Regressão Logística é a análise discriminante linear. A técnica de análise discriminante tem como base a suposição de que existem duas populações de indivíduos, distinguidas por “0” e “1”, cada uma caracterizada por uma distribuição multivariada de um conjunto de atributos,  $x$ , incluindo fatores como idade, renda, tamanho da família, histórico de crédito, ocupação, entre outros. Um indivíduo com atributo vetor  $x_i$ , é traçado a partir de uma das duas populações, e deseja-se que determine qual. A análise é realizada através de um pedido a atribuição de um ‘score  $Z$ ’, calculado como:  $Z_i = a + b'x_i$  (Greene, 1998)

Resumindo, num contexto atual, usa-se este método ‘score  $Z$ ’ para classificar os mutuários como bons pagadores ( $D = 0$ ) ou maus pagadores ( $D = 1$ ). As aplicações de indivíduos designados para a classe  $D = 1$  são rejeitados (Maddala, 1983 *apud* Greene, 1998)<sup>37</sup>.

Podemos ainda referir através de Purohit, Mahadevan e Kulkarni (2012) que o objetivo de um modelo de regressão logística está na obtenção de uma equação de regressão, que possa prever qual de dois ou mais grupos deve ser classificado como bom ou mau crédito. A regressão logística também prevê a probabilidade de um alvo binário ou ordinal adquirir o

---

<sup>36</sup> Schreiner, M. (1999). A scoring model of the risk of costly arrears at a microfinance lender in Bolivia.

<sup>37</sup> Maddala, G., 1983. Limited Dependent and Qualitative Variables in Econometrics, Cambridge University Press, New York.

evento de interesse (por exemplo, o pagamento do crédito ou *credit default*). Através de uma função com uma ou mais variáveis independentes (ou seja, o montante de crédito, o trabalho do mutuário, o propósito do crédito). O modelo *logit* é representado pela função logística resposta P (Y) da forma:

$$P(y) = 1/(1 + \exp(-z)) \text{ onde } z = b_0 + \sum b_i x_i \quad \forall i = tom \quad (2.4)$$

A função P (Y) descreve uma variável dependente y contendo dois ou mais resultados qualitativos. Z é a função de variáveis X, n independente chamado preditores, e b representa os parâmetros. As variáveis X podem ser variáveis categóricas ou contínuas de qualquer distribuição. O valor de P (Y) que varia de 0 a 1 indica a probabilidade de que uma variável dependente y pertence a um dos dois ou mais grupos. O diretor de máxima verossimilhança pode ser comumente usada para calcular estimativas dos parâmetros b.

Segundo Ong, Huang e Tzeng (2005) os modelos de regressão logística são um dos instrumentos estatísticos mais utilizados nos problemas de classificação. O Modelo de regressão logística, ao contrário de outras ferramentas estatísticas (como exemplo, análise discriminante ou a regressão linear ordinária), pode acompanhar vários tipos de funções de distribuição, tais como Gamble, Poisson, normal, etc., por isso torna-se o método mais utilizado.

#### 2.7.4 Análise Discriminante

Para iniciar uma análise discriminante terão de ser cumpridas várias etapas, nomeadamente, i) remoção de *outliers*; ii) efetuar testes de normalidade às variáveis; iii) Homogeneidade de matrizes; e, por fim, iv) Existência de diferenças significativas entre as variáveis.

Respetivamente, a primeira etapa trata da remoção de variáveis que causam grande impacto nas médias, e que provocam um aumento das variâncias transmitindo um resultado errado da significância estatística. Deste modo, os *outliers* deverão ser removidos antes do início da análise. Nos testes de normalidade às variáveis assume-se que os dados representam uma distribuição normal multivariada. Na última etapa são utilizadas medidas estatísticas para medição de possíveis diferenças significativas.

A análise discriminante é uma técnica estatística multivariada utilizada em inúmeros problemas que envolvem a separação de duas ‘populações’ e a inclusão de novas observações em grupos diferentes.

Segundo Reis (2001:205), «na análise de regressão é pressuposto que a variável dependente tem distribuição normal enquanto as variáveis independentes são determinísticas. Na análise

discriminante o pressuposto reverte-se: pressupõe-se que as variáveis independentes têm distribuição conjunta normal multivariada enquanto a dependente é fixa e de tipo nominal (assume os valores de 1, 2,...,  $k$ , consoante o número de grupos).»

Pelo mesmo autor, «o objetivo da análise discriminante é encontrar uma combinação linear das variáveis independentes que minimize a probabilidade de incorreta classificação dos indivíduos.»

A análise discriminante é levado a cabo através de mais do que uma combinação linear das variáveis independentes utilizadas ( $X_i$ ). Cada combinação linear ( $Y_i$ ) constitui uma **função discriminante** (ibid.)

$Y_i = a_{i0} + a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p$ , em que:

$a_{ij}$  = Coeficientes de ponderação

$X_j$  = variáveis discriminantes não-normalizadas.

Os pesos na fórmula são derivados com estatísticas, mas a matemática é a parte fácil, o difícil é coletar dados sobre o desempenho e características de empréstimos passados, a transportar pontuação no processo de empréstimos na avaliação atual, e ainda ajustar as organizações a aceitar uma técnica tão fundamental e bem-sucedida até ao momento atual (Schreiner, 2000). Segundo Abdou, El-Masry e Pointon (2007), o modelo de análise discriminante assume especificamente que i) as variáveis de previsão são medidos em uma escala de intervalo; ii) as matrizes de covariância das variáveis de previsão são iguais para os dois grupos; e, por último, iii) as variáveis de previsão seguem uma distribuição normal multivariada.

### 2.7.5 Modelos de Seleção das variáveis

Para a seleção automática no SPSS, adição ou remoção de variáveis, destacam-se dois métodos, o *Forward* que consiste Método de seleção *Stepwise*, este procedimento inicia-se com um modelo que não contenha variáveis explicativas. A ideia do método é adicionar uma variável de cada vez, selecionando em primeiro lugar aquela que apresentar um valor de correlação mais elevado, em módulo, com a variável resposta, e assim consequentemente, até que o processo pára quando o aumento do coeficiente de determinação, devido à inclusão de uma nova variável explicativa no modelo não é mais importante. Por outra, tem-se o *Backward*, que enquanto o método *Forward* começa sem nenhuma variável no modelo e adiciona variáveis a cada passo, o *Backward* faz o oposto. Este incorpora inicialmente todas as variáveis, e ao longo do processo cada uma pode ou não ser eliminada. A primeira variável a ser removida

é aquela que apresenta um menor coeficiente de correlação parcial com a variável resposta. (www.portalaction.com.br) O método *Stepwise* é um dos métodos mais utilizados e consiste na combinação dos dois métodos anteriores (*Forward e Backward*). Este, inicia com uma variável (a que apresentar maior correlação com a variável resposta), e a cada passo do *Forward*, depois de incluir uma variável, aplica o *Backward* para ver se será descartada alguma variável. Continua-se o processo até este não incluir ou excluir nenhuma variável.

### CAPÍTULO III – Metodologia

É a fase do estudo onde são exibidos os dados que foram utilizados para tratamento e de que forma foram transformados/trabalhados.

Foram utilizadas formas distintas de obtenção dos dados, através de uma pesquisa no âmbito qualitativo e quantitativo.

No que respeita à parte qualitativa ou exploratória, é assim designada pelo facto de ser um assunto ainda em fase embrionária, precisamente na entidade financeira base de estudo.

Respetivamente à parte quantitativa, significa que todos os valores da pesquisa podem ser transportados numericamente, ou seja, são mensuráveis sendo por isso classificados e analisados.

Para a execução da presente pesquisa foram exportados de uma base de dados vários indicadores de clientes distintos, com o objetivo de constituir uma amostra sólida, sendo o alvo da análise. O local onde foram efetuadas essas exportações para estudo, foi uma instituição financeira, atualmente localizada em Portugal, conseguindo exportar um número de casos considerados suficientes para obtenção dos resultados pretendidos.

Os clientes têm características variadas/mistas, não respeitando a qualquer tipo de perfil previamente elaborado.

A primeira fase dos dados para a elaboração do estudo foram exportados em Novembro/2014, e correspondem a um período de 4 anos de propostas (2011 – 2014) concedidas de crédito hipotecário no banco em estudo.

Os dados, para serem trabalhados, tiveram de ser submetidos a uma mineração, de modo a que os elementos dos clientes estivessem ‘limpos’ para introdução no programa estatístico, SPSS. Desta forma, estão criadas as condições necessárias para o início da análise e construção da base do modelo de *scoring*.

De acordo com o que foi descrito anteriormente, o primeiro ponto foi separar o crédito habitação dos outros tipos de créditos, ou seja, juntar todas as operações de financiamento imobiliário. Desse modo, depreende-se assim que a base de dados à qual nos cingimos refere-se aos produtos de crédito habitação do banco concedidos em Portugal, no regime geral de crédito à habitação.

Os produtos são habitação própria, troca de banco, habitação própria (oferta Sénior), troca de casa, habitação para arrendamento e terrenos. A finalidade destes produtos baseia-se na aquisição do imóvel, aquisição e obras, construção e por último, somente obras.

Através de várias ferramentas conseguimos distinguir as operações morosas das não morosas, colocando um ponto de partida no qual se definem as operações em incumprimento (todas as que o mutuário está em atraso no pagamento à mais de 30 dias).

Para analisar o perfil de endividamento de um cliente, tem-se como apoio a Central de Responsabilidades de Crédito (CRC).

Através da CRC consegue-se verificar todos os créditos que o cliente já possui na banca e quanto será o encargo gerado que irá pagar mensalmente pela totalidade das suas responsabilidades creditícias.

Na análise de risco de crédito, é um apoio imprescindível, principalmente na análise da proposta de crédito quando existem dúvidas sobre a concessão ou não do crédito.

A CRC é um sistema de informação, gerido pelo Banco de Portugal, constituído por informação recebida das entidades participantes sobre responsabilidades efetivas ou potenciais decorrentes das operações de crédito e por um conjunto de serviços relativos ao seu processamento e difusão.

A CRC tem como principal objetivo apoiar as entidades participantes na avaliação do risco da concessão de crédito. Para o efeito, estas entidades acedem à informação agregada das responsabilidades de crédito de cada cliente no conjunto do sistema financeiro, podendo ainda aceder a informação relativa a potenciais clientes na sequência de um pedido de concessão de crédito ou mediante autorização dos mesmos.

A CRC contém informação sobre responsabilidades de crédito efetivas assumidas por qualquer particular, empresa ou outra entidade perante as entidades participantes, independentemente de se encontrarem em situação regular (informação positiva) ou em incumprimento (informação negativa); bem como sobre responsabilidades de crédito potenciais que representem compromissos irrevogáveis no sistema financeiro. Na grande maioria dos casos, a informação contida na CRC corresponde a informação positiva, comprovando a capacidade de pagamento e a pontualidade do cliente de crédito na data a que se refere essa mesma informação. Por esse motivo, a CRC não configura qualquer espécie de 'lista negra' (Caderno do Banco de Portugal, nº 5).

### **3.1 Análise Descritiva**

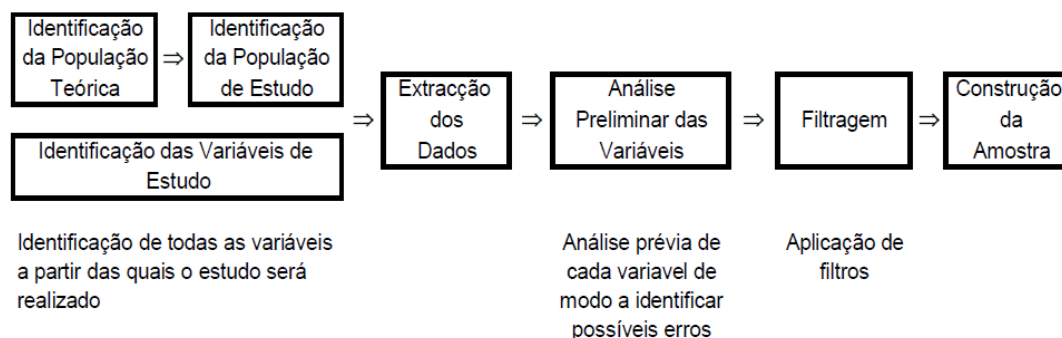
A amostra para o estudo efetuado consistiu num conjunto de 493 propostas de crédito. Apenas foram selecionadas as propostas de crédito vivo, ou seja, o conjunto das aprovadas que ainda podem ser analisadas do lado do risco enquadrados no Regime Geral de Crédito à Habitação.

Especificamente, 399 em situação normal e 94 casos em situação de incumprimento.

### 3.2 Recolha dos dados

O processo da recolha dos dados e construção da amostra pode ser resumido na seguinte figura:

**Figura 3.1** - Processo de recolha de dados



**Fonte:** Adaptado de Fazenda (2008:29)

A precisão de um sistema de pontuação de crédito irá depender do cuidado com que se desenvolve. Os dados em que o sistema se baseia, necessitam de ser uma rica amostra de ambos os empréstimos com bom e mau desempenho.

Conforme Gool, Baesens, Sercu e Verbeke (2009) a preparação de dados trata da escolha e da criação das variáveis desejadas e tratamento de valores e valores atípicos. Não existe uma abordagem universalmente aceite para escolher as variáveis explicativas dos candidatos para um modelo de pontuação de crédito perfeito. A revisão da literatura indica que a maioria dos autores se referem a consultoria especializada e (ou) aos estudos prévios para explicar as suas opções.

Os dados devem ser atualizados, e os modelos devem ser frequentemente avaliados para garantir que as mudanças nas relações entre potenciais fatores de desempenho e empréstimo sejam capturados (Mester, 1997).

As agências de crédito são donas de uma “ampla gama continuamente atualizada de informações detalhadas sobre milhões de créditos pessoais históricos dos consumidores.” [tradução livre do autor]. Os registos de crédito são usados para gerar relatórios financeiros, que são um produto do núcleo de agências financeiras e o ingrediente fundamental na pontuação de crédito. Chamamos a isso de ‘dados’ - que são totalmente reportados às agências de crédito nacional (Robinson, 2014).

A mineração de dados refere-se à extração de variáveis/indicadores úteis a partir de uma grande amostra. O processo de mineração de dados passa pela identificação das características necessárias, e com o uso de técnicas específicas para análise dos mesmos pretende-se alcançar resultados marcantes para a tomada de decisões estratégicas (Berry e Linoff, 2004 *apud* Yap, Ong e Husain, 2011)<sup>38</sup>.

### 3.2.1 Validação dos Dados em Excel

Para controlar o tipo de dados e valores introduzidos nas células teve de se proceder à validação dos dados como 1º passo da iniciação do trabalho prático.

Este passo permite definir restrições aos dados recolhidos, configuração esta que prevê a implementação de dados inválidos. Para a execução da validação de dados, o excel possui o grupo de 'Ferramenta de Dados' com essa finalidade.

A validação dos dados é um passo de extrema importância quando se pretende que a pesquisa seja dada a vários leitores, de modo a que os resultados sejam consistentes e fidedignos. Esta ferramenta pode ser utilizada na restrição de dados a itens predefinidos numa lista, tal como a criação de uma própria lista com base num intervalo desejado; e restrição de números fora de um intervalo especificado; restrição de dados ou horas fora de determinado período de tempo; limitar o número de caracteres de texto e, por fim, validar os dados com base em fórmulas ou valores noutras células ([www.support.office.com](http://www.support.office.com)).

### 3.3 Variáveis

Para a implementação de um modelo de *Credit Scoring* terão de se seleccionar as variáveis necessárias para a correlação com o modelo.

Segundo Fazenda (2008) existem as variáveis da operação que caracterizam a mesma; as variáveis de comportamento da operação que são as que determinam se as operações estão ou já estiveram em incumprimento; as variáveis sociodemográficas que caracterizam os interveientes num contexto social e demográfico, e por fim, as variáveis de relação com a instituição, que caracterizam a relação do cliente com o banco nos 12 meses anteriores à operação. Conforme Woldridge (2002), as variáveis serão dependentes e independentes, as primeiras respetivamente, poderão usar uma regressão múltipla para explicar a dependente binária, com

---

<sup>38</sup> Berry, M. J. A., & Linoff, G. S. (2004). *Data mining techniques: For marketing, sales, and customer support*. New York: John Wiley and Sons, Inc.

valores zero ou um; a segunda, por último, são variáveis binárias ou dicotômicas onde agregam informação qualitativa em modelos de regressão estatística.

A análise dos dados relevantes da amostra inclui uma comparação dos valores reais e teóricos de características demográficas dos clientes do banco, e a estimativa de coeficientes da análise discriminante (Mylonakis e Diacogiannis, 2010).

Tendo em conta a exploração bibliográfica que tem vindo a ser efetuada, denota-se que as variáveis mais utilizadas no processo de criação de modelos de apoio à decisão para crédito habitação são:

- Idade;
- Número de elementos do agregado familiar;
- Zona de residência;
- Habilitações Literárias;
- Vínculo/Situação Profissional;
- Estado Civil.

Estas variáveis dizem respeito ao perfil do cliente, são variáveis sociodemográficas já referidas anteriormente. Ainda assim, existe outro grupo de variáveis que definem a relação do cliente com o banco e que caracterizam a componente quantitativa da operação, considerando estas como sendo de extrema importância na concessão de financiamento imobiliário.

As variáveis são:

- Valor da Prestação;
- Prazo do empréstimo;
- Taxa de Esforço;
- Rácio Loan-to-Value;
- Valor do Imóvel;
- Rendimento Anual Declarado;
- Rendimento mensal do proponente
- Garantia/Hipoteca.

De acordo com Nguyen (2014) variáveis como estado civil, género, educação, emprego e o tempo de residência na morada atual são variáveis utilizadas por autores como Orgler (1971), Steenackers e Goovarts (1989), Banasik et al., (2003), Chen e Huang (2003), Sarlija et al., (2004), Hand et al., (2005), Sustersic et al., (2009).

Variáveis como antiguidade no emprego, rendimentos, empréstimos com outros bancos, montante do empréstimo solicitado, duração do empréstimo, finalidade do empréstimo e

garantias são outro grupo de variáveis utilizadas por um leque de investigadores no âmbito dos modelos de *scoring* de crédito, como Greene (1998), Schreiner (2004), Ong et al., (2005), Kleimeier e Dinh (2007), Kocenda e Vojtek (2011).

No entanto, existem outros autores que relacionam estes dois grupos, como Hand e Henley (1996), M. B. Yobas et al., (2000), Jacobson e Roszbach (2003), Abdou, Pointon e El-Masry (2007), Yap, Ong e Hussain (2011), Samreen e Zaidi (2012), Itoo e Selvarasu (2015), entre muitos outros.

Schreiner (1999) defende justamente que a maioria dos modelos de *scoring* de crédito possuem uma lista reduzida de variáveis. O mesmo refere que a maioria dos estudos consiste em variáveis no âmbito sociocultural e ainda, económica, referentes ao perfil do mutuário e do empréstimo, respetivamente. O mesmo diz ainda que (1999:5),

*The econometric model uses knowledge of the traits of past borrowers at the time of disbursement and of their subsequent repayment performance to infer future repayment risk for potential borrowers whose traits are known and who have passed the standard evaluation.*

Existem no entanto variáveis elaboradas consoante a falta das mesmas, e na medida em que são necessárias para dar respostas significativas. Ao contrário da constante de variáveis semelhantes na maioria dos estudos, Kocenda e Vojtek (2011) utilizam os dados do cliente para a construção de rácios considerando assim um maior número de variáveis para realização de testes. Exemplo dessas variáveis são, como os mesmos denominam o *Credit Ratio 1* e o *Credit Ratio 2* em que o primeiro respetivamente refere-se à percentagem do rendimento que é gasta em despesas, e o segundo refere-se à proporção do rendimento disponível do cliente para o salário mínimo oficial válido no momento do pedido do empréstimo.

Como será de fácil perceção, existem inúmeras variáveis que poderiam estar incluídas no presente estudo. Contudo, serão incluídas apenas as consideradas relevantes através da bibliografia consultada e das operações de crédito realizadas na instituição financeira.

A finalidade deste estudo passa por identificar quais as variáveis que têm um forte impacto no incumprimento, particularmente, no crédito à habitação.

Objetivamente, o estudo visa mostrar o impacto que várias condicionantes têm no incumprimento, fatores sociais (perfil do proponente) como fatores naturalmente intrínsecos à operação, sendo LTV, o valor de avaliação do imóvel, montante a financiar, taxa de esforço,

etc. A execução do modelo de *Credit Scoring* terá a junção de todas as variáveis que efetivamente influenciam a probabilidade de incumprimento/*default*.

Para um melhor entendimento do que será realizado nos pontos seguintes, denotou-se a importância de mostrar o conjunto de atividades a desenvolver no crédito à habitação.

### **3.3.1 Atividades desenvolvidas no processo de financiamento imobiliário**

O crédito hipotecário tem três fases de execução sendo composto pela primeira por pré-aprovação, onde é efetuada uma simulação ao cliente. Nesta fase é apresentada uma proposta provisória, o cliente decide se aceita ou não o exposto, caso este aceite a proposta é enviada para o departamento de análise de risco de crédito do Banco em análise (DARC).

Na DARC, passará a uma fase de aprovação em que se irão verificar todos os elementos carregados na proposta de crédito e analisar os indicadores de maior relevância neste tipo de financiamento, como o LTV e a taxa de esforço. São naturalmente analisados outros elementos, como os encargos globais que o cliente possui, as responsabilidades efetivas e potenciais (como explicado anteriormente através da CRC do BdP), são também efetuadas avaliações ao imóvel para que não haja diferenças do crédito solicitado pelo cliente e o valor do mesmo. Após a análise efetuada pelo analista de risco, o banco pode aprovar ou não o crédito, ou ainda poderá aprovar mediante algumas condições em caso de dúvida do cliente conseguir pagar ou não o encargo proposto.

Numa fase final, o cliente é avisado da decisão tomada pela instituição financeira, da concessão ou não do crédito. Caso não tenha sido concedido, o gestor poderá constituir nova proposta ao cliente e se o mesmo aceitar novas condições a proposta voltará para a DARC com a finalidade de ser reapreciada.

Segundo Gestel e Baesel (2009), há características importantes que definem um crédito, como exemplo do que a generalidade dos bancos tem tido em conta, temos:

Pessoa Física: medidas do carácter dos mutuários e integridade (por exemplo, a reputação, honestidade, etc.)

Capital: mede a diferença entre os ativos (por exemplo, carro, casa, etc.) e passivos (por exemplo, as despesas com rendas) do devedor;

Colateral: mede o tipo de garantias prestadas (por exemplo, casa, carro, fiadores, etc.)

Capacidade: mede a capacidade do devedor para pagar o crédito (por exemplo, o cargo que possui no emprego, as rendas que paga, os rendimentos que obtém, etc.)

Condição: medidas circunstanciais inerentes ao mutuário (por exemplo, condições de mercado, pressão competitiva de caráter sazonal).

O indicador mais importante é o tempo com o banco, seguido por sexo, número de empréstimos, e duração do empréstimo. Segundo investigação efetuada por Than e Kleimeier (2006).

### **3.4 Método utilizado no Estudo**

Com o aumento do incumprimento, torna-se necessário a seleção de um método que nos dê de forma mais rápida a previsão de possíveis incumprimentos, ou seja, a previsão de comportamentos de risco.

Há vários métodos constantemente utilizados nesta área de estudo, sendo os mais utilizados a análise discriminante, *logit*, *probit*, redes neuronais, indução de regras e árvores de decisão, algoritmos genéricos, conjuntos aproximados, etc.

O objetivo do estudo de caso, não passa por comparar os modelos mais evidenciados, mas sim, a utilização de um que nos dê as respostas certas de forma mais simples e prática.

A análise Discriminante e a Regressão Logística estão inseridas nos métodos estatísticos multivariados de dependência, pelo facto de relacionarem um conjunto de variáveis independentes com uma variável dependente (Sharma, 1996). Para os resultados de uma análise discriminante serem considerados verdadeiros tem de se supor a existência de uma normalidade multivariada. Isto significa que variáveis independentes terão de corresponder a alguns critérios e a uma distribuição normal.

Uma das vantagens da regressão logística é que exige um menor número de pressupostos, contrariamente à análise discriminante (SPSS, 2010).

Para além do mencionado anteriormente, tem-se que o resultado obtido de uma análise discriminante corresponde a uma pontuação, não transportando o elemento probabilístico, levando com uma maior frequência ao uso do Modelo Logístico (Ohlson, 1980).

#### **3.4.1 Modelo de Regressão Logística - *Logit***

Tendo em conta o objetivo do presente trabalho, deduziu-se que o modelo mais adequado seria o Modelo de Regressão Logística.

Segundo Nguyen (2014) a Regressão Logística continua a ser o método mais popular, onde atualmente, mesmo nas maiores instituições financeiras do mundo, esta é sempre considerada como uma das principais abordagens para prever situações de *default*.

Segundo Gool, Baensens, Sercu e Verbeke (2009), dois modelos foram desenvolvidos para analisar a pontuação de crédito a pequenos credores: um modelo *logit* binário com base em

variáveis binárias codificadas (3.1), e um modelo *logit* binário com base no peso da evidência variáveis codificadas (3.2).

$$\pi_i = E(Y_i = 1) = \frac{1}{1 + e^{-\text{logit}_i}}, \quad (3.1)$$

$$\text{logit}_i = \beta_0 + \beta_{11}x_{i11} + \beta_{12}x_{i12} + \dots + \beta_{148}x_{i148}$$

$Y_i$  = Variável dependente binária;

$B_0$  = Interceção;  $\beta_i$  = coeficiente de regressão;

$X_i$  = Variável explicativa *dummy*

Este modelo (1) é baseado numa regressão logística, com base numa classificação entre bom ou mau da entrada binária como variável dependente.

$$\pi_i = E(Y_i = 1) = \frac{1}{1 + e^{-\text{logit}_i}}, \quad (3.2)$$

$$\text{logit}_i = \beta_0 + \beta_1x_{i1} + \beta_2x_{i2} + \dots + \beta_{16}x_{i16}$$

$Y_i$  = Variável dependente binária;

$B_0$  = Interceção;  $\beta_i$  = coeficiente de regressão;

$X_i$  = peso da evidência codificada da variável explicativa.

### 3.4.2 Análise Discriminante

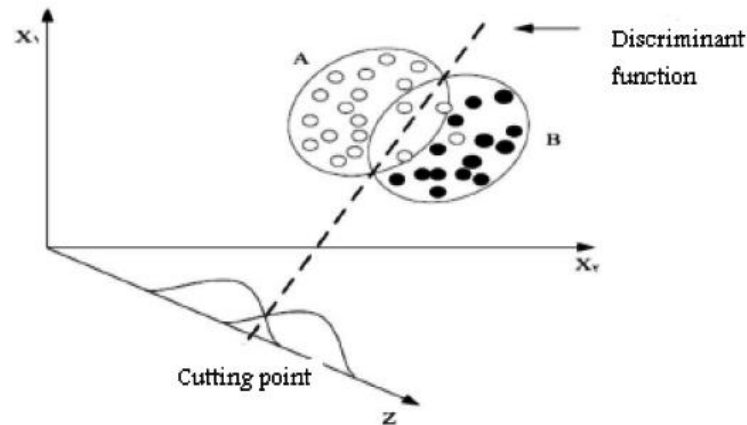
Segundo Pardoe, Yin e Cook (2006) a análise discriminante é utilizada para categorizar os valores de uma variável dependente nominal com dois ou mais aspetos.

Na verdade, para os casos em que as variáveis dependentes nominais e variáveis independentes são quantitativas, a análise discriminante é utilizada para prever as alterações feitas nas variáveis dependentes (associações de grupo) com base no uso dessas variáveis. Neste método, as variáveis dependentes, também chamadas de variáveis de previsão são combinadas para formar um novo leque de variáveis e é atribuída uma pontuação discriminante a cada caso. Esta nova variável chamada função discriminante é calculada de uma forma que os clientes são classificados em diferentes classes da variável dependente (critério) com base nos valores/pontuação obtida.

A função discriminante é, na prática, a forma de encontrar uma combinação linear das variáveis independentes padronizadas com a criação da variância máxima entre os grupos/classes

em que a variância intra-grupo atinge o nível mínimo. Para tal acontecimento, a função discriminante é criada com base nos grupos centróides (meio), de modo que haja o mínimo de sobreposição entre os grupos.

**Figura 3.2 - Análise Discriminante**



Fonte: Dastoori e Mansouri (2013: 29)

Para explicar o que foi mencionado acima, em termos matemáticos,  $X_1, X_2, \dots, X_n$  é suposto ser variáveis independentes e  $Z$  é um multinível (categórica) variável dependente. A análise discriminante visa determinar uma função linear da seguinte forma:

$$Z_i = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (3.3)$$

De modo que a probabilidade seguinte seria em um nível máximo:

$$P(Z = z | (X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)) \quad (3.4)$$

Como pode ser visto na Figura (3.2), quando a variável dependente é composta por dois níveis (grupos), destina-se a atribuição de novas observações ( $X_1, X_2, \dots, X_n$ ) a um destes dois grupos com base em  $Z$  (a função discriminante). Para fazer isso, o coeficiente  $\beta$  representa a participação de cada variável na função de pontuação, escolhido de forma que a pontuação  $Z$  resultante das funções acima, discrimine, de forma otimizada, os grupos. Além disso, o valor de  $Z_i$  é calculado de uma forma que os intervalos entre os meios (centróide)

está no nível máximo nos dois grupos ou seja, ZA e AB. Estas exigências serão satisfeitas se a função de  $\beta$  é determinada da seguinte forma:

$$\beta = \Sigma^{-1}(XA - XB) \quad (3.5)$$

Em que:

XA - é o vetor de média de  $n$  variáveis nos casos em estudo no Grupo A;

XB - é o vetor de média de  $n$  variáveis nos casos em estudo no Grupo B;

$\Sigma$  - A matriz de variância-covariância para  $n$  variáveis independentes.

### **3.5 Ferramentas utilizadas – SPSS**

Para realização do modelo proposto ter-se-á de recorrer a ferramentas especializadas, nomeadamente, o SPSS. Segundo Maroco (2003) o SPSS é o sistema mais utilizado no âmbito das ciências sociais. Este programa é produzido pela SPSS, Inc. Uma companhia americana originária de Chicago, formada especificamente para produzir programas de análise estatística para essa matéria. Desde 1994, o SPSS está representado em Portugal pela PSE (Produtos e Serviços de Estatística, Lda – [www.pse.pt](http://www.pse.pt)).

De acordo com o mesmo autor, no SPSS a análise dos dados faz-se em 4 etapas, sendo elas:

- 1) Introduzir as variáveis da amostra no editor de dados;
- 2) Selecionar o procedimento de análise de um dos menus;
- 3) Selecionar as variáveis a analisar;
- 4) Examinar e interpretar os resultados.

## CAPÍTULO IV – Estudo Empírico

### 4.1 Análise Descritiva da Amostra

Tendo concluído o capítulo anterior, a amostra encontra-se preparada para a execução de uma primeira análise. Inicialmente a investigação centrou-se numa análise univariada de modo a obter o escrutínio sintetizado do conjunto de dados, ou seja, do universo da amostra entre operações de incumprimento e não incumprimento.

A amostra está caracterizada por um total de 493 clientes com operações de crédito em curso, dos quais 399 estão em situação regular e 94 em situações de incumprimento. De salientar, que os casos que se encontram em morosidade correspondem ao não pagamento das responsabilidades financeiras de crédito num período superior a 30 dias.

**Tabela 4.1** - Classificação Total da Amostra

	Universo de Operações (Análise <i>Default</i> )		
	Classificação	Nr. Operações	Percentagem
Cumprimento	0	399	80,9
Incumprimento	1	94	19,1
<b>Total</b>	<b>-</b>	<b>493</b>	<b>100</b>

Do quadro acima representado, podemos concluir que o número de operações de crédito em incumprimento é muito menor que as operações em situação normal. Assim, no total da amostra (100%) visualiza-se um grau de incumprimento de 19,1% e, créditos em situação regular a cifrar-se nos 80,9%.

### 4.2 Análise Descritiva das Variáveis

O número total das variáveis independentes, para efeitos de análise, corresponde a um número de 18 fatores, sociodemográficos e económicos. Nos primeiros, respetivamente, estão incluídas 9 variáveis, sendo *Age; Education Level; Relashion Bank; Divorce; Big city; Labor Bond; Nr. Of dependents; Job; Other Loans*.

O Segundo grupo de variáveis, de carácter económico, são as restantes 9 variáveis, sendo *o LTV; Loan to Income; Income; Loan Amount; Outgoings Amount; Purpose; Loan Duration; Loan Provision; Guarantee*; seguidamente, irá ser apresentado um resumo de todas as variáveis incluídas na amostra.

### 4.2.1 Idade (*Age*)

Através do quadro apresentado abaixo podemos definir a variável ‘idade’ para as operações de crédito na sua globalidade. Esta variável foi utilizada em inúmeros estudos, nomeadamente, por autores como Hand e Henley (1996), Greene (1998), Gool, Baensens, Sercu e Verbeke (2000), Jacobson e Roszbach (2003), Schreiner (2003), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Leung, Cheong e Cheong (2008), Yap, Ong e Husain (2011), Kocenda e Vojtek (2011), Samreen e Zaidi (2012), Kinda e Achonu (2012), Nguyen (2014), Itoo e Selvarasu (2015), entre outros. Visualiza-se assim que pela sua utilidade terá importância na inclusão do modelo.

**Tabela 4.2** - Descrição da Variável – Idade (*AGE*)

	Amostra	Idade Mínima	Idade Máxima	Média	Desvio Padrão
<b><i>AGE</i></b>	493	24,5	70,1	39,551	9,5392

Da totalidade da amostra, pode-se concluir que a idade mínima de pedidos de Crédito à Habitação regista os 24-25 anos de idade. A idade máxima apresenta-se nos 70 anos de idade, cujo valor está em concordância com a Idade Limite de Amortização Real (ILAR) do empréstimo, mencionada no Regulamento de Crédito à Habitação, a praticar pela instituição financeira em estudo. Depreende-se ainda, que a média de idade dos clientes da amostra encontra-se entre os 39 e 40 anos com uma margem (desvio) de 9 anos, ou seja, entre os 30 e os 49 anos são as idades onde se regista um maior volume de solicitações de crédito hipotecário.

Analisando a variável independente com a variável dependente, pode-se concluir que dentro das operações normais ( $y=0$ ), a idade mínima de solicitação de crédito é de 21 anos, e a máxima é de 67 anos. A média da idade nas operações em situação regular é de 38 anos, com um desvio (margem) de 8 anos.

No universo das operações morosas ( $y=1$ ), observa-se a idade mínima de 25 anos e máxima de 70 anos (Idade Limite). A média de idade nas operações morosas é de 46 anos, com um desvio (margem) de 11 anos.

### 4.2.2 Educação (*Education Level*)

Variável utilizada em estudos por autores como Schreiner (1999), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Coravos (2010), Kocenda e Vojtek (2011), Samreen e

Zaidi (2012) entre outros. Através do quadro abaixo representado, pode-se visualizar a variável de ‘nível educação’.

**Tabela 4.3** - Descrição da variável – Nível de Educação (*EDUCATION LEVEL*)

	<i>EDUCATION LEVEL</i>		
	Classificação	Nr. Operações	Percentagem
Outros estudos inferiores	0	374	75,9
Ensino Superior	1	119	24,1
<b>Total</b>	-	<b>493</b>	<b>100</b>

Denota-se assim que para candidatos com curso superior (Licenciatura, Mestrado ou grau superior) atribuiu-se o valor 1, enquanto para casos em que os estudos sejam de nível inferior à classificação de 1, passarão a estar assinalados com o valor 0.

Pelo exposto, conclui-se que a maior parte dos clientes da amostra (374 clientes) têm apenas o 12º ano ou inferior, e que uma pequena parte (119 clientes) possui estudos superiores.

#### 4.2.3 Relação com o Banco (*Relashion Bank*)

Para descrever esta variável, autores como Viganó (1993), Hand e Henley (1996), Yobas, Crook e Ross (1999), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Samreen e Zaidi (2012), entre outros, utilizaram uma classificação como 0 os candidatos que não tinham qualquer relação com a instituição financeira, e contrariamente, foram classificados com o valor 1 as operações em que os proponentes já tiveram ou têm relação com o banco. Pelo quadro representado abaixo, visualiza-se assim, que o maior pedido de crédito é solicitado por clientes que já possuem relações creditícias com o banco. Sendo que da totalidade da amostra verifica-se que o diferencial entre cliente e não cliente não se indica relevante.

**Tabela 4.4** - Descrição da Variável – Relação Bancária (*RELASHION BANK*)

	<i>RELASHION BANK</i>		
	Classificação	Nr. Operações	Percentagem
Não é cliente	0	220	44,6
É cliente	1	273	55,4
<b>Total</b>	-	<b>493</b>	<b>100</b>

#### 4.2.4 Estado Civil (*Divorce*)

Autores como Viganó (1993), Hand e Henley (1996), Schreiner (1999), Jacobson e Roszbach (2003), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Yap, Ong e Husain (2011), Kocenda e Vojtek (2011), Samreen e Zaidi (2012), Nguyen (2014), Itoo e Selvarasu (2015), entre outros, utilizaram esta variável em estudos com as mesmas características.

A variável '*Divorce*' apresenta-se classificada como 1 para os casos em que o proponente está legalmente divorciado, e como contrapartida, para os casos em que estão solteiros, em união de facto ou casados ou outro estado atribuiu-se o valor de 0.

**Tabela 4.5** - Descrição da Variável – Estado Civil (*DIVORCE*)

	<i>DIVORCE</i>		
	Classificação	Nr. Operações	Percentagem
Outros	0	451	91,5
Divorciado	1	42	8,5
<b>Total</b>	<b>-</b>	<b>493</b>	<b>100</b>

Denota-se, que há um número reduzido de casos de divórcio. A constituição da amostra revela uma percentagem de 8,5% de divorciados em 100%.

#### 4.2.5 Outros empréstimos (*Other Loans*)

Pelo quadro 4.6, visualiza-se que da totalidade da amostra, a maior parte dos clientes já possui outros créditos para além do crédito à habitação a que se está a propor.

Esta variável foi utilizada por autores como Jacobson e Roszbach (2003), Kraft, Kroisandt e Muller (2003), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Kocenda e Vojtek (2011), Samreen e Zaidi (2012), Nguyen (2014), entre outros.

**Tabela 4.6** - Descrição da Variável - Outros empréstimos (*OTHER LOANS*)

	<i>OTHER LOANS</i>		
	Classificação	Nr. Operações	Percentagem
Sem créditos	0	149	30,2
Possui outros créditos	1	344	69,8
<b>Total</b>	<b>-</b>	<b>493</b>	<b>100</b>

#### 4.2.6 Naturalidade do proponente (*Bigcity*)

Variável utilizada por autores como Greene (1998), Jacobson e Roszbach (2003), Than e Kleimeier (2006), Leung, Cheong e Cheong (2008), Yap, Ong e Husain (2011), Nguyen (2014), entre outros.

Através do quadro apresentado, a variável encontra-se indicada por 1 os clientes que são naturais de Lisboa, e por 0 os que não são naturais da capital.

**Tabela 4.7** - Descrição da Variável – Naturalidade (*BIGCITY*)

	<i>BIGCITY</i>		
	Classificação	Nr. Operações	Porcentagem
Outros	0	446	90,5
Naturalidade Lisboa	1	47	9,5
<b>Total</b>	<b>-</b>	<b>493</b>	<b>100</b>

#### 4.2.7 Regime Laboral (*Labor Bond*)

Autores como Hand e Henley (1996), Greene (1998), Gool, Baesens, Sercu e Verbeke (2000), Than e Kleimeier (2006), Itoo e Selvarasu (2015), entre outros, utilizaram esta variável para chegar ao objetivo proposto com o presente trabalho.

Esta variável obteve a classificação para 1 os clientes que se encontram em regime de trabalho efetivo, e 0 outros tipos (incluindo estágios, contratos sem termo, entre outras).

**Tabela 4.8** - Descrição da Variável – Regime Laboral (*LABOR BOND*)

	<i>LABOR BOND</i>		
	Classificação	Nr. Operações	Porcentagem
Outros	0	79	16
Efetivo	1	414	84
<b>Total</b>	<b>-</b>	<b>493</b>	<b>100</b>

Depreende-se, pelo diferencial da amostra, que na solicitação do crédito à habitação, talvez por ser um dos pedidos mais longos a que um cliente se propõe, o maior número de pedidos de crédito os proponentes encontram-se em regime de efetividade, o que transporta um maior conforto às operações.

#### 4.2.8 Número de Dependentes (*Nr. Dependents*)

Esta variável foi regularmente utilizada por Greene (1998), Than e Kleimeier (2006), Yap, Ong e Husain (2011), Samreen e Zaidi (2012), entre outros.

Na classificação desta variável, denotou-se que o número máximo de dependentes na amostra era de 3. Os casos em que possuímos o valor de '-1' quer dizer que não há informação sobre esse proponente, o que poderá corresponder a casos em que possuem mais de 3 filhos, ou simplesmente sem dependentes. Para melhor detalhe desta variável, segue-se o quadro abaixo.

**Tabela 4.9** - Descrição da Variável – Nr. de dependentes (*Nr. OF DEPENDENTS*)

	<b><i>NR. OF DEPENDENTS</i></b>		
	Classificação	Nr. Operações	Porcentagem
-1	Sem informação	239	48,5
0	-	134	27,2
1	-	76	15,4
2	-	37	7,5
3	-	7	1,4
<b>Total</b>	-	<b>493</b>	<b>100</b>

Conclui-se que esta variável não se considera a mais apropriada para fornecer as respostas que pretendemos, uma vez que o maior número de operações não possui informação definida.

#### 4.2.9 Atividade Profissional (*Job*)

Para a classificação desta variável, a cada profissão corresponde um código, como tal, na análise profunda da variável obteve-se uma enorme variedade de resultados para cada proposta/cliente. Deste modo, apenas serão tidas em conta as profissões que englobam um maior número de casos, ou seja, as que sejam de maior relevância e que representem um maior número na amostra.

Pelo exposto, em destaque segue a profissão como empresário onde se regista o maior número de casos, traduzindo 49 clientes exercendo essa atividade. Seguidamente, estão as profissões como motorista e operário fabril, com 25 casos cada. Com 20 casos apresentam-se os clientes como sócios gerentes. Decrescendo, 14 clientes com a profissão de gerente e 14

clientes como auxiliar administrativo. Com menor destaque, seguem 12 vendedores e 12 empregados de escritório.

Esta variável foi utilizada em diversos estudos por autores como Greene (1998), Gool, Baensens, Sercu e Verbeke (2000), Than e Kleimeier (2006), entre outros.

#### 4.2.10 Rácio Financiamento/Garantia (*Loan-To-Value*)

Autores como Ohlson (1980), Avery, Bostic, Calem e Canner (1995), Gurný e Gurný (2013) entre outros, foram de muitos os que utilizaram esta variável para modelos de *Scoring* de Crédito.

Através do quadro representado abaixo, pode-se visualizar a variável LTV na globalidade da amostra.

**Tabela 4.10** - Descrição da Variável - Rácio F/G (*LOAN TO VALUE*)

	Amostra	LTV mínimo	LTV Máximo	Média	Desvio Padrão
<b>LTV</b>	493	10,00%	100,50%	59%	20%

Da totalidade da amostra, pode-se concluir que o LTV mínimo de pedidos de Crédito à Habitação regista os 10%. O LTV máximo apresenta-se nos 100%. Depreende-se ainda que a média do LTV da amostra encontra-se nos 59% com uma margem (desvio) de 20%, ou seja, entre os 40% e os 80%, aproximadamente, são as percentagens de LTV onde se regista um maior volume de concessões de crédito hipotecário.

#### 4.2.11 Taxa de Esforço (*Loan-To-Income*)

Através do quadro representado abaixo, pode-se visualizar a variável TE na globalidade da amostra.

A taxa de esforço representa o quociente entre a prestação proposta do empréstimo e o rendimento (disponível) líquido mensal do proponente.

Ohlson (1980), Avery, Bostic, Calem e Canner (1995), Bellotti e Crook (2007), Gurný e Gurný (2013), são os vários autores que utilizaram nas suas pesquisas esta variável.

**Tabela 4.11** - Descrição da Variável - Taxa de esforço (*LOAN TO INCOME*)

	Amostra	Mínimo	Máximo	Média	Desvio Padrão
<b>LOAN TO INCOME</b>	493	0,034	2,65	0,3467	0,3098

Com o quadro apresentado pode-se verificar que a TE mínima da amostra está na casa dos 3%, o que significa que o proponente possui rendimentos muito superiores ao encargo que está a ser proposto, minimizando o risco de incumprimento.

O máximo registado é de 2,65, significando que o proponente está a solicitar um empréstimo 2 vezes superior aos seus rendimentos. A média de TE contratada é de 35%, o que ainda assim, se encontra acima do admissível. A margem está na casa dos 31%.

#### **4.2.12 Rendimento Anual Líquido Declarado (*Income*)**

No que respeita à variável rendimento, autores como Avery, Bostic, Calem e Canner (1995), Jacobson e Roszbach (2003), Than e Kleimeier (2006) utilizaram esta variável. Muitos autores, na sua generalidade utilizaram esta variável, devido à sua extrema importância na análise de crédito a particulares.

As variáveis relacionadas com os rendimentos estão desdobradas em várias categorias, uma vez que são tidos em conta diferentes critérios para cada uma delas, individualmente. Deste modo, tem-se:

- Rendimento Anual Líquido Declarado (*Income*)
- Rendimento Mensal Líquido Declarado (RM\_IRS)
- Rendimentos Anuais Através do valor líquido dos Recibos (*Monthly Net Income*)
- Rendimento Médio Líquido dos recibos (RML\_Recibos)

Para efeitos de apreciação de crédito os valores que são tidos em conta são sempre os valores líquidos. Após uma realização da estatística descritiva, podemos verificar que o valor médio do rendimento líquido declarado dos candidatos inseridos na análise é de 26.029,5€ e o máximo é de 3.048.584€.

Face ao rendimento mensal líquido para efeito declarado em IRS, o valor máximo cifra-se em 133.490€ com média de 1.660€ mensais.

Para o rendimento anual proveniente dos recibos, tem-se um máximo de 170.800€ e uma média de 15.586€.

No rendimento médio líquido dos recibos, tem-se um máximo de 14.233€ e uma média de 1.299€.

A única variável tida em linha de conta para o estudo, foi apenas a *INCOME*, sendo que é a única variável suficiente para estimar os parâmetros pretendidos. Na utilização de outras variáveis relacionadas, estaríamos presente uma situação de multicolinearidade, onde haveria uma sobreposição de variáveis.

#### 4.2.13 Valor do Empréstimo (*Loan Amount*)

Através da análise descritiva consegue-se visualizar através do quadro abaixo o montante do empréstimo.

Autores como Jacobson e Roszbach (2003), Abdou, Masry e Pointon (2007), Nguyen (2014) entre outros, utilizaram esta variável para pesquisa em trabalhos semelhantes.

**Tabela 4.12** - Descrição da Variável – Valor do Empréstimo (*LOAN AMOUNT*)

	Amostra	LA mínimo	LA Máximo	Média
<b><i>LOAN AMOUNT</i></b>	493	10.000 €	2.500.000 €	91.630 €

Denota-se assim que o valor mínimo solicitado para o crédito à habitação é de 10.000€ e o máximo é de 2.500.000€. Valor mínimo que coincide com o estipulado no Regulamento de Crédito à Habitação da entidade bancária, em que o mínimo solicitado permitido é de 10.000€.

#### 4.2.14 Despesas (*Outgoings*)

Através do quadro representado abaixo, pode-se visualizar a variável *Outgoings* na globalidade da amostra. Conforme Yobas, Crook e Ross (1999), esta foi também uma das variáveis a ter em conta para efeitos de concessão de crédito.

**Tabela 4.13** - Descrição da Variável – Despesas (*OUTGOINGS*)

	Amostra	Mínimo	Máximo	Média
<b><i>OUTGOINGS</i></b>	493	0 €	25.065 €	430,00 €

Assim, depreende-se que o máximo de despesas que o cliente possui é de 25.065€ anuais, com uma média de 430€.

#### 4.2.15 Propósito do Empréstimo (*Purpose*)

Autores como Gool, Baesens, Sercu e Verbeke (2000), Kraft, Kroisandt e Muller (2003), Kinda e Achonu (2012) utilizaram esta variável para estudo do modelo. Esta variável, está sempre presente nas propostas de crédito à habitação, uma vez que a finalidade do crédito está obrigatoriamente evidenciada em todas as operações.

Através do mesmo método utilizado anteriormente, denota-se que o maior número de pedidos de crédito são para aquisição de habitação própria, e em segundo, destaca-se naturalmente a finalidade para a construção.

**Tabela 4.14 - Descrição da Variável – Propósito do Empréstimo (*PURPOSE*)**

	<b><i>PURPOSE</i></b>
	Nr. Operações
Aquisição	305
Construção	81
Consolidação de CH	43
Obras	30
Multifinalidades	10
Trf. De Crédito	10
Aquisição e Obras	9
Outras	5
<b>Total</b>	<b>493</b>

#### 4.2.16 Duração do Empréstimo (*Loan Duration*)

Than e Kleimeier (2006), Abdou, Masry e Pointon (2007) utilizaram esta variável nas suas investigações.

Através do quadro apresentado abaixo podemos definir a variável do prazo do empréstimo para as operações de crédito na sua globalidade.

**Tabela 4.15 - Descrição da Variável – Duração do Empréstimo (*LOAN DURATION*)**

	Amostra	Prazo mínimo	Prazo máximo	Média	Desvio Padrão
<b><i>LOAN DURATION</i></b>	493	0	50	27,99	10,44

Da totalidade da amostra, pode-se concluir que o prazo mínimo de pedidos de Crédito à Habitação é nulo, em casos de tornas ou de casos em que se adquire parte de um imóvel, em que as situações são imediatas, não sendo um pagamento longo. O prazo máximo apresenta-

se nos 50 anos, cujo valor está em concordância com a ILAR do empréstimo, mencionada no Regulamento de Crédito à Habitação, a praticar pela instituição financeira em estudo. Depreende-se ainda que o prazo médio dos empréstimos da amostra encontra-se entre os 27 e 28 anos com uma margem (desvio) de 10 anos, ou seja, entre 18 e 38 anos é o prazo de maior frequência nas contratações de crédito hipotecário.

#### 4.2.17 Prestação do Empréstimo (*Loan Provision*)

Ohlson (1980), Bellotti e Crook (2007) foram dos muitos autores que utilizaram esta variável nos seus testes de modelos de *Credit Scoring*. No entanto, esta variável é mais uma das que está implicitamente presente, pois consta obrigatoriamente em todas as propostas de crédito à habitação solicitadas à instituição bancária. Através do quadro representado abaixo, pode-se visualizar a variável *Loan Provision* na globalidade da amostra.

**Tabela 4.16** - Descrição da Variável – Prestação do empréstimo (*LOAN PROVISION*)

	Amostra	Mínimo	Máximo	Média
<b><i>LOAN PROVISION</i></b>	493	0	24.348 €	510,95 €

Assim, depreende-se que a prestação máxima que o cliente suporta é de 24.348€ anuais, com uma média de 510,95€, correspondente à que maioritariamente se pratica, na ordem dos 500€.

#### 4.2.18 Garantia (*Guarantee*)

Utilizada por Schreiner (2003), Jacobson e Roszbach (2003), Abdou, Masry e Pointon (2007), Kinda e Achonu (2012), entre outros, a Garantia foi mais uma das variáveis utilizadas no estudo. Através do quadro representado abaixo, pode-se visualizar a variável *Guarantee* na globalidade das propostas de crédito à habitação.

**Tabela 4.17** - Descrição da Variável – Garantia (*GUARANTEE*)

	<b><i>GUARANTEE</i></b>		
	Classificação	Nr. Operações	Porcentagem
2ª Hipoteca e outras	0	25	5,10%
1ª Hipoteca	1	468	94,90%
<b>Total</b>	-	<b>493</b>	<b>100%</b>

Denota-se que a maioria das propostas de crédito está garantida pela 1ª hipoteca (garantia real), sendo que uma pequena parte está garantida por uma 2ª hipoteca ou outro tipo de garantia. De salientar que este tipo de garantia não é muito usado, estando o quadro em concordância com o que se pratica regularmente, sendo praticamente 100% das propostas de crédito colateralizada pela 1ª hipoteca.

### 4.3 Modelo de Regressão

Após o descritivo e a caracterização de cada variável, o método agora a ser utilizado será uma análise multivariada, ou seja, um modelo de escolha binária. Como foi explicado em capítulos anteriores, será agora utilizado o modelo de Regressão Logística. Como também foi evidenciado, existem outros modelos que poderiam ser utilizados aquando de um problema semelhante ao exposto, como será exemplo a análise discriminante. Não obstante, o modelo escolhido centra-se na sua simplicidade e facilidade de interpretação tendo presente os dados que disponibilizados e os resultados esperados.

Seguidamente, será descrito o *output* utilizado no programa SPSS após a realização da análise estatística do modelo de Regressão Logística. Tratando-se de Regressão Logística, os erros seguem uma distribuição binomial tendo como base o teste de razão de verosimilhança. Na base das decisões de inserção e remoção de variáveis o método utilizado foi o **'Enter'**, seleção automática. Este método baseia-se na entrada de todas as variáveis em bloco de uma única vez. Um dos pontos de relevo num trabalho de regressão logística é a descrição do ponto de corte. Ponto que reflete a proporção existente entre as duas situações da amostra total, as operações normais e em incumprimento.

No estudo presente, houve a necessidade da remoção de alguns dados para consolidação da amostra, casos em que alguns dados não estavam em concordância com os demais. Desta forma, a proporção não é real, pelo que nestas situações o ponto de corte utilizado é de 0,5 definindo assim probabilidades iguais para as duas amostras, situações regulares e incumprimento.

Em estudos semelhantes, autores como Viganó (1993), Hand e Henley (1996), Schreiner (1999), Gool, Baensens, Sercu e Verbeke (2000), Than e Kleimeier (2006), Abdou, Masry e Pointon (2007), Bellotti e Crook (2007), Coravos (2010), Yap, Ong e Husain (2011), Kocenda e Wojtek (2011), Kinda e Achonu (2012), Samreen e Zaidi (2012), Gurný e Gurný (2013), Nguyen (2014) e Robinson (2014), entre outros, foram dos muitos a utilizar o modelo *Logit* pela sua simplicidade e fácil interpretação dos resultados.

### 4.3.1 O Procedimento no SPSS

Inicialmente, o procedimento tido no SPSS (versão 20) é composto pela distribuição da amostra. Deste quadro representado abaixo, adaptado do *Output* do SPSS, verifica-se que a amostra contém 493 dados. A inexistência de ‘casos em falta’ deve-se ao facto deste tipo de dados já terem sido eliminados antecipadamente.

**Tabela 4.18 - Case Processing Summary**

Casos não ponderados		Nº	Percentagem
Casos Seleccionados	Incluídos na análise	493	100
	Casos em Falta ( <i>missing</i> )	0	.0
	Total	493	100.0
Casos não Seleccionados		0	.0
Total		493	100.0

Houve ainda a necessidade de alteração da correspondência da variável dependente, onde o quadro abaixo representa essa mudança em SPSS.

**Tabela 4.19 - Dependent Variable Encoding**

Original Value	Internal Value
0	0
1	1

Consecutivamente, apresenta-se a tabela de classificação, onde o modelo é considerado unicamente com a constante. Este facto explica-se através de uma situação hipotética, em que se considera que todos os créditos à habitação provenientes na amostra estejam em situação regular, significando assim pelo resultado evidenciado no quadro, que a taxa de acerto seria de 80,9%.

**Tabela 4.20** - Tabela de Classificação a,b

Observações			Previsão		
			DEFAULT		Percentagem correta
			0	1	
Step 0	DEFAULT	0	399	0	100.0
		1	94	0	.0
Percentagem geral					80.9

a. Constant is included in the model.

b. The cut value is .500

De seguida, pelo *output* SPSS, obteve-se o quadro da introdução das variáveis. No estudo em particular pelo método **Enter**.

**Tabela 4.21** - Omnibus Tests of Models Coefficients

		Qui-Quadrado	Grau de Liberdade (Df)	Significância
Step 1	Passo (Step)	99.733	4	.000
	Bloco (Block)	99.733	4	.000
	Modelo (Model)	99.733	4	.000

Através do quadro 4.21, depreende-se que as variáveis foram seleccionadas em apenas um passo, obtendo-se o modelo final esperado. Consta-se que o coeficiente é significativo no passo efetuado.

Na observação do 'Modelo (*Model*)' observa-se que o modelo é bastante significativo (*Chi-Square* de 99.733), excluindo assim a premissa de que as variáveis independentes não explicam a variável dependente (variável Y que representa a dualidade da amostra entre operações em situação regular e não regular).

De seguida, é apresentado a adaptação ao modelo geral. Serão expostas 3 medidas de análise à adaptação geral do modelo.

**Tabela 4.22** - Model Summary

Step	-2 Log likelihood	Cox & Snell R <sup>2</sup>	Nagelkerke R <sup>2</sup>
1	380.639 <sup>a</sup>	.183	.294

<sup>a</sup> Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

O Cox & Snell não poderá exceder o valor de 1, visto não obter um valor que atinja esse limite. O critério Nagelkerke surge para corrigir os limites do critério anterior. Quanto maior forem estes dois parâmetros melhor será o modelo final.

Tendo-se realizado apenas 1 passo, pela maior significância deduz-se que o modelo de Regressão Logística apresenta elevada significância.

Posteriormente, realizada a avaliação do ajuste geral do modelo, tem-se de seguida a Validação do Modelo. No quadro abaixo representado encontra-se espelhada finalmente, a classificação.

**Tabela 4.23 - Table Classification a**

Observações			Previsão		
			DEFAULT		Percentagem correta
			0	1	
Step 1	DEFAULT	0	392	7	98.2
		1	66	28	29.8
Percentagem geral					85.2

a. The cut value is 0,5.

Interpretando os valores evidenciados no quadro 4.23, pode-se dizer que o modelo apresenta uma melhor explicação no caso de operações em situação normal, do que nas operações em incumprimento. Pelo valor retirado do quadro, explica 98,2% das operações normais contra 29,8% das operações em incumprimento, ou seja, numericamente explica 392 operações de crédito das 399 em situação normal, e apenas 28 das 94 operações em incumprimento.

De seguida, o quadro 4.24 representado, irá mostrar as variáveis finais, e respetivos Betas (parâmetros) obtidas através da Regressão Logística. Tem-se ainda para análise a estatística *Wald* que traduz o peso que cada variável independente possui, individualmente na explicação da variável dependente. Tendo em consideração as variáveis incluídas no quadro abaixo, conclui-se que a significância foi admissível.

**Tabela 4.24 - Variables in the Equation**

		$\beta$	S.E	Wald	df	Sig.	Exp (B)
Step 1 <sup>a</sup>	AGE	.068	.014	24.077	1	.000	1.070
	LTV	2.139	.679	9.908	1	.002	8.488
	LOAN TO IN- COME	1.901	.496	14.696	1	.000	6.691
	RELASHION BANK	.651	.294	4.893	1	.027	1.918
	Constant	-6.712	.806	69.428	1	.000	.001

a. Variable(s) entered on step 1: AGE, LTV, LOANTOINCOME, and RELASHIONBANK.

Como se pôde visualizar, para obtenção do modelo final foram selecionadas 4 variáveis das 18 incluídas no modelo. O modelo final encontra-se, assim, composto pelas variáveis *Age*, *LTV*, *Loan to Income* e *Relashion Bank*. Resumindo, estas foram as 4 variáveis que explicaram da melhor forma o comportamento da variável dependente, ou seja, apresentam maior significância.

A significância corresponde a uma regra de decisão que permite a decisão de rejeitar ou não uma hipótese estatística com base nos resultados da amostra.

Através do método *Enter* utilizado na seleção das variáveis, obtiveram-se as que têm grau de significância elevada (evidenciadas na tabela 4.24), as restantes da amostra foram naturalmente excluídas por não terem nível de significância estatística.

O nível de significância é o limite que se toma como base para afirmar que um certo desvio é decorrente do acaso ou não.

São aceites como estatisticamente significativos os níveis de  $p= 0,05$  e  $p= 0,01$ , ou seja, 5% e 1%.

Na prática, considera-se satisfatório o limite de 5% de probabilidade de erro, não sendo significativas as diferenças que tiverem uma probabilidade acima desse limite.

O nível de significância deve ser estabelecido antes do experimento ser realizado e corresponde ao risco que se corre de rejeitar uma hipótese verdadeira ou aceitar uma hipótese falsa (Beiguelman, 1996).

Finalmente, após os parâmetros estimados pelo modelo em estudo obteve-se a seguinte equação final:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i = X_i \beta + u_i$$

Em que  $Y_i$  é a variável dependente que toma valores de 0 (Normal) e 1 (Incumprimento).  
Especificamente, 1 se  $Y \geq 0$  e 0 se  $Y < 0$ .

Pelo exposto, tem-se:

$$Y_i = -6,712 + (0,068 \times \text{Age}) + (2,139 \times \text{LTV}) + (1,901 \times \text{LoanToIncome}) + (0,651 \times \text{RelationBank})$$

#### 4.3.2 Diagnóstico de Multicolinearidade

Torna-se importante medir a correlação das variáveis explicativas na presença de variáveis regressoras. Situações em que não haja qualquer relação linear designam-se de ortogonais, situação que dificilmente se verifica.

Atribui-se o termo colinearidade quando a análise de correlação está presente entre duas variáveis. Do mesmo modo, atribui-se o termo multicolinearidade quando a análise de correlação é entre mais que duas variáveis.

Um dos pontos mais importantes na interpretação e análise dos modelos de regressão logística são as estimativas dos seus coeficientes.

Uma das formas práticas de diagnosticar a Multicolinearidade será através do *Variance Inflation Factor* (VIF).

A título indicativo, um  $VIF > 10$  indica problemas de multicolinearidade.

No SPSS, ao efetuar o ajuste do modelo através do teste de Multicolinearidade obteve-se o seguinte quadro abaixo, representado do *Output*:

**Tabela 4.25 - Coeficientes a**

MODEL	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	-0,52	,085		-6,127	0		
AGE	,010	,002	,236	5,515	0	,870	1,149
LTV	,284	,079	,146	3,586	0	,966	1,035
RELASHION-BANK	,070	,033	,089	2,098	,036	,895	1,117
LOANTOINCOME	,342	,055	,269	6,246	0	,858	1,166

a. Variável dependente: *DEFAULT*

O VIF é dado por:

$$VIF_j = \frac{1}{1 - R_j^2}.$$

O  $VIF_j$  mede o quanto a variância do coeficiente *Beta* é inflacionada pela sua colinearidade. Como já referido anteriormente, denota-se que os valores de VIF encontram-se ligeiramente acima de 1, o que se depreende que a Multicolinearidade é fraca ([www.portalaction.com](http://www.portalaction.com)).

## CAPÍTULO V – Conclusão

De modo geral, considera-se que o objetivo do trabalho foi conseguido, uma vez que as tarefas executadas ao longo do trabalho vão de acordo às necessidades da instituição.

Esse objetivo passou por desenvolver um modelo de *Credit Scoring* para o crédito à habitação. Hoje em dia torna-se fulcral possuir sistemas e modelos desenvolvidos capazes de dar respostas rápidas aos clientes, visto que a concorrência cada vez é mais agressiva, não permitindo falhas.

Atualmente, a generalidade do crédito, e não somente do crédito à habitação, apresentam períodos de reservas face a outras épocas. Contudo, não descurando da importância do crédito em geral, o crédito habitação apresenta algumas características que o torna demasiado importante para uma instituição financeira. Características como o prazo do empréstimo, e o montante traduzem-se em ganhos para a IC, fidelizando ainda os clientes através de *Cross-Selling*. O objetivo deste trabalho, passou exatamente por definir as variáveis que são preponderantes em situações de incumprimento. Para este fim, teve-se como material de suporte uma base de dados de 2011 – 2014 (inclusive) disponibilizada pela instituição financeira alvo de estudo. Essa base de dados continha informação de vários indicadores afetos ao cliente e à qualidade da proposta em si. No entanto, ainda que os dados tenham sido extraídos diretamente da fonte que se pretendia investigar, tiveram de se assumir algumas limitações.

*Enter* foi o método utilizado para inclusão das variáveis no SPSS. Incluíram-se todas as variáveis da amostra, total de 18 variáveis, com o objetivo de serem encontradas as de maior significância na previsão de incumprimento, tendo-se obtido as seguintes: *AGE*, *LOAN TO INCOME*, *LTV* e *RELASHIONBANK*. As restantes variáveis não foram incluídas no modelo por não terem significância estatística.

O método utilizado na realização do estudo de caso, foi o método de Regressão Logística – *Logit*, o qual segundo a literatura investigada é dos modelos de maior facilidade de aplicação e interpretação de resultados.

### 5.1 Limitações e Sugestões para Trabalhos Futuros

As sugestões futuras vêm em linha com as limitações. Sugere-se, em trabalhos que venham a ser desenvolvidos, que se detenham bases de dados sólidas, com condições apropriadas para serem alvo de estudo.

Denotou-se a inexistência de dados relativos a determinados clientes, campos que não foram preenchidos ao nível da formalização da proposta em *Workflow* de Crédito Hipotecário (WFCH) e carregados incorretamente.

Pelo referido, houve a impossibilidade de mostrar o peso de determinadas variáveis, que alguns investigadores tinham como relevantes.

Acredita-se que quanto mais variáveis forem incluídas, melhores resultados se poderão chegar, por isso, para trabalhos futuros deverão ser considerados o maior número de variáveis.

Outras limitações a incluir são a dimensão da amostra, período de 2011-2014 que para efeitos de estudo considera-se reduzida para tirar conclusões fortes; e, aponta-se ainda o facto de o estudo incidir apenas sobre uma instituição financeira, o que torna as conclusões mais específicas e centralizadas apenas na mesma.

Com efeito, estudos desta natureza deverão abranger uma amostra de maior dimensão, nomeadamente, englobando várias instituições financeiras.

Outra forma de alargar o trabalho desenvolvido será a aplicação do modelo não só a bancos nacionais, como também Europeus, numa óptica de alargar também a carteira de clientes e futuramente, pudermos obter uma maior e melhor dimensão de resultados.

Outra limitação encontrada na aplicação deste modelo está sobretudo relacionada com a proporção das operações morosas da população. Como tal, foi utilizado o ponto de corte 0,5, assumindo-se probabilidades idênticas para os 2 grupos da amostra (morsas e não morsas) e ainda a falta de outros parâmetros que expliquem casos de incumprimento.

Em sùmula, futuramente torna-se fundamental, que para estudos desta natureza sejam obtidas base de dados cada vez mais completas e, para isso, parte da instituição financeira reforçar essa necessidade junto dos seus colaboradores, apelando que para a implementação de metodologias perfeitas necessita-se da obtenção de informação, igualmente, perfeita.

Tendo salientes todas as limitações, conseguiu-se de grosso modo, um modelo final coerente e ajustado com elevado nível de precisão, de 85,2%.

## Referências Bibliográficas

- ABDOU, H. & POINTON, J. (2011) **Credit scoring, statistical techniques and evaluation criteria: a review of the literature**, *Intelligent Systems in Accounting, Finance & Management*, 18 (2-3), pp. 59-88.
- ABDOU, H., POINTON, J. & EL-MASRY, A. (2007) **On the applicability of credit scoring models in Egyptian Banks**, *Banks and Bank Systems/ Volume 2, Issue 1*.
- ALTMAN & SAUNDERS (1998). **Credit Risk Measurement: Developments over the last 20 Years**. *Journal of Banking & Finance*, 21(11-12), pp. 1721-1742.
- AVERY, ROBERT B., RAPHAEL W. BOSTIC, P. S. CALEM, & G. B. CANNER. **Credit Risk, Credit Scoring and the Performance of Home Mortgages**, *Federal Reserve Bulletin* 82 (July 1996), pp. 621- 48.
- BANCO DE PORTUGAL: **Boletim Estatístico**. [em linha] (Dezembro de 2014) [Consultado em Março de 2015]. Disponível em [https://www.bportugal.pt/pt-PT/EstudosEconomicos/Publicacoes/BoletimEconomico/Publicacoes/Bol\\_Econ\\_dezembro\\_p.pdf](https://www.bportugal.pt/pt-PT/EstudosEconomicos/Publicacoes/BoletimEconomico/Publicacoes/Bol_Econ_dezembro_p.pdf)
- BANCO DE PORTUGAL: **Relatório de Estabilidade Financeira**. [em linha] (Nov. 2014) [Consultado em Março 2015]. Disponível em [https://www.bportugal.pt/pt-PT/EstabilidadeFinanceira/Publicacoes/RelatoriodeEstabilidadeFinanceira/Publicacoes/Relatorio%20de%20Estabilidade%20Financeira\\_nov2014.pdf](https://www.bportugal.pt/pt-PT/EstabilidadeFinanceira/Publicacoes/RelatoriodeEstabilidadeFinanceira/Publicacoes/Relatorio%20de%20Estabilidade%20Financeira_nov2014.pdf)
- BANCO DE PORTUGAL: **Central de Responsabilidades de Crédito**. Caderno nº5 do BdP [em linha] (Maio de 2011). [Consultado em Março de 2015]. Disponível em <https://www.bportugal.pt/pt-PT/PublicacoesIntervencoes/Banco/CadernosdoBanco/Biblioteca%20de%20Tumbnails/Central%20de%20Responsabilidades%20de%20Cr%C3%A9dito.pdf>
- BANCO DE PORTUGAL: **Modelo de Avaliação de Riscos**. Departamento de Supervisão Bancária. [em linha] (2007). [Consultado em Março de 2015]. Disponível em [https://www.bportugal.pt/pt-PT/OBancoeEurosisistema/ConsultasPublicas/Documents/Consulta\\_BP\\_2\\_07\\_MAR.pdf](https://www.bportugal.pt/pt-PT/OBancoeEurosisistema/ConsultasPublicas/Documents/Consulta_BP_2_07_MAR.pdf)
- BANCO DE PORTUGAL: **Situação Patrimonial dos setores não financeiros**. [em linha] (Nov. 2013) [Consultado em Março/2015] Disponível em: [https://www.bportugal.pt/pt-PT/EstudosEconomicos/Publicacoes/RelatorioEstabilidadeFinanceira/Publicacoes/Cap1.2\\_p.pdf](https://www.bportugal.pt/pt-PT/EstudosEconomicos/Publicacoes/RelatorioEstabilidadeFinanceira/Publicacoes/Cap1.2_p.pdf)

- BANCO DE PORTUGAL (2011): **Comunicado Relativo a aviso do Banco de Portugal sobre reforço do rácio “Core Tier 1” das Instituições de Crédito.** O banco e o Eurosistema. [em linha] (Abril, 2011) [Consultado em Março 2015]. Disponível em <https://www.bportugal.pt/pt/PT/OBancoeoEurosistema/ComunicadoseNotasdeInformacao/Paginas/comb20110407.aspx>
- BANK FOR INTERNATIONAL SETTLEMENTS (2012): **Results of the Basel III Monitoring exercise as of 30 June 2011.** Switzerland. ISBN web: 92-9197-128-6
- BARREIRA, A., F., (2012): **A análise dos Modos de Falha e seus Efeitos aplicada a um processo do setor Bancário.** Dissertação de Mestrado. Universidade Nova de Lisboa.
- BASTO, Rita B.; (2013). **Uma Política Macroprudencial para a Estabilidade Financeira.** Artigos Banco de Portugal. 1-16
- BEIGUELMAN, B. **Curso Prático de Bioestatística** – 4ª ed. (1996) [consultado em Junho de 2015]. Disponível em <http://www.ufpa.br/dicas/biome/biotestes.htm>
- BELLOTTI T. & CROOK J., **Credit scoring with macroeconomic variables using survival analysis,** *Journal of the Operational Research Society*, vol. 60, pp. 1699-1707, 2008
- CHAKRAVARTY S., JHA A. N., (2012) **Viability of “Credit Scoring in Microfinance” for Developing Countries.** *International Review of Social Sciences and Humanities*. Vol. 3, No. 1, pp. 104-107
- CHRISTENSEN, R. (1997) **Log-Linear Models and Logistic Regression.** Springer Texts in Statistics. ISBN 0-387-98247-7
- DINH, T. H. T., & KLEIMEIER, S. (2007). **A credit scoring model for Vietnam’s retail banking market,** *International Review of Financial Analysis*, 16, 471-495
- DZELIHODZIC A., DONKO D., (2013) **Data Mining Techniques for Credit Risk Assessment Task.** *International Burch University*
- EUROPEAN BANKING AUTHORITY: **Parecer da Autoridade Bancária Europeia sobre Boas Práticas para a Concessão Responsável de Crédito Hipotecário.**
- FAZENDA, N. (2008): **Determinantes do Default no Crédito Habitação Hipotecário.** Dissertação de Mestrado, Instituto Superior de Economia e Gestão, Universidade Técnica de Lisboa.

- FRADE, C. (2003): **Desemprego e Sobreendividamento dos Consumidores – Contornos de uma Ligação Perigosa**. Projeto de Desemprego e Endividamento das Famílias - Relatório Final. Centro de Estudos Sociais da Faculdade de Economia de Coimbra (CES).
- FINANCIAL STABILITY BOARD (2012): **Principles for sound residential Mortgage Underwriting Practices**. 1-13
- GESTEL, T. V., & BAESENS, B. (2009): **Credit Risk Management**. Oxford University Press. New York. ISBN 978-0-19-954511-7
- GOOL, J. V., BAESENS, B., SERCU, P. & VERBEKE, W. (2009) **An analysis of the applicability of credit scoring for microfinance**.
- GREENE W., (1998) **Sample Selection in Credit Scoring Models**, *Japan and the World Economy*, 10, 299-316
- HAND D. J. & HENLEY W. E. **Statistical Classification Methods in Consumer Credit Scoring: a Review**. *J. R. Statist. Soc. A 160, Part 3, pp. 523-541*
- HECKMAN, J., 1979. **Sample selection bias as a specification error**. *Econometrica*, 47
- Inquérito à Situação Financeira das Famílias** (2012): Banco de Portugal e Instituto Nacional de Estatística
- ITOO, R. A. & SELVARASU A. (2015) **Case study of consumer credit scoring: A proposal**. *International Journal of Recent Scientific Research. Vol. 6, Issue, 4, pp.3627-3461*
- JACOBSON T., ROSZBACH K. (2003) **Bank Lending Policy, Credit Scoring and value-at-risk**. *Journal of Banking & Finance*, 27, 615–633.
- JORNAL Público – **Bancos vão voltar a apostar no crédito habitação mas os jovens não**. [em linha] (Maio, 2014) [Consultado dia 20 Jan. 2015] Disponível em: <http://www.publico.pt/economia/noticia/bancos-va-voltar-a-apostar-no-credito-a-habitacao-mas-os-jovens-nao-1635032>
- KOCENDA E. & VOJTEK M. (2011) **Default Predictors in retail credit scoring: Evidence from Czech banking data**. *The William Davidson Institute, Working paper number 1015*
- LAUREANO, M., & BOTELHO, M. **SPSS: o meu manual de consulta rápida**. Lisboa. Edições Sílabo, 2010.

- LORETTA, J. M. (1997), **What's the point of Credit Scoring**, *Business Review*, Federal Reserve Bank of Philadelphia
- MAROCO, J. (2003). **Análise estatística - com utilização do SPSS**. Lisboa: Edições Sílabo, 2ª edição.
- MARTIN, D., 1977. **Early warning of bank failure: A logit regression approach**. *Journal of Banking and Finance* 249-276.
- MYLONAKIS, J. & DIACOGIANNIS G. (2010) **Evaluating the Likelihood of using linear discriminant analysis as a commercial bank card owners credit scoring model**. *International Business Research Vol. 3, No. 2*.
- MULLER M., & HARDLE W. (2002) **Exploring Credit Data**. *Berlin, Germany*
- MURTEIRA, Bento; RIBEIRO, Carlos Silva; ANDRADE E SILVA, João; PIMENTA, Carlos – **Introdução à Estatística**. Lisboa: Escolar Editora, 2010. ISBN 978-072-592-282-8
- NEVES, João Carvalho – **Análise e Relato Financeiro: Uma Visão Integrada de Gestão**. Lisboa: Texto Editores, 2012. ISBN 978-972-47-4326-4
- NEW YORK TIMES: **Need a loan? Ask the computer: Credit Scoring changes small – business lending** [em linha] (Abril, 1995) [Consultado dia 15 de Março. 2015] Disponível em: <http://www.nytimes.com/1995/04/18/business/company-need-loan-ask-computer-credit-scoring-changes-small-business-lending.html>
- NGUYEN H. (2014) **How is credit scoring used to predict default in China? How is credit scoring used to predict default in China?** *Université de Paris Ouest Nanterre La Défense*.
- OHLSON J. A., (1980) **Financial Ratios and the Probabilistic Prediction of Bankruptcy**. *Journal of Accounting Research*, Vol. 18, No. 1, pp. 109-131
- ONG C., HUANG J. & TZENG G. (2005) **Building credit scoring models using genetic programming**. *Expert Systems with Applications xx, 1-7*
- PARDOE I., YIN X. & COOK R. D. (2006) **Graphical tools for quadratic discriminant analysis**.
- PEREIRA, Luís N.; CHORÃO, Luís, R. (2007). **Identificação das Características dos Clientes Associadas ao Risco de Crédito**. *Encontros Científicos*, 12-26.

- PESTANA, M. & GAGEIRO, J. **Descobrimo a Regressão – com a complementariedade do SPSS**. Lisboa: Edições Sílabo, 2005. ISBN: 972-618-394-4.
- PUROHIT, S. U. MAHADEVAN V. & KULKARNI A. N. (2012) **Credit evolution model of loan proposals for Indian Banks**. *International Journal of Modeling and Optimization, Vol. 2, No. 4*.
- REIS, E., (2001): **Estatística Multivariada Aplicada**. Edições Sílabo, Lisboa. ISBN 972-618-247-6
- REIS, Elizabeth; MELO, Paulo; ANDRADE, Rosa; CALAPEZ, Teresa – **Estatística Aplicada**. Lisboa: Edições Sílabo, 1996. ISBN 972-618-132-1
- ROBINSON, YU (2014) **Knowing the score: New data, underwriting and marketing in the consumer credit marketplace**. *A Guide for Financial Inclusion Stakeholders*
- ROSENFELD M. J., (2002) **A Brief Orientation, or Where Log-Linear Models Fit in to the big picture**.
- SAMREEN A., & ZAIDI F. B., (2012) **Design and Development of Credit Scoring Model for the Commercial banks of Pakistan: Forecasting Creditworthiness of Individual Borrowers**. *International Journal of Business and Social Science*. Vol. 3. No. 17
- SANTOS, A.; TELES, N., SERRA, N.; (2014): **Finança e Habitação em Portugal**. Cadernos do Observatório. Centro de Estudos Sociais da Universidade de Coimbra.
- SEMEDO, D; (2009). **Credit Scoring: Aplicação da Regressão Logística vs Redes Neurais Artificiais na Avaliação do Risco de Crédito no mercado Cabo-Verdiano**. Dissertação de Mestrado, Universidade Nova de Lisboa.
- SOUSA, Maria José; BAPTISTA, Cristina Sales – **Como fazer Investigação, Dissertações, Teses e Relatórios Segundo Bolonha**. 4ª ed. Lisboa: Pactor, 2011. ISBN 978-989-693-001-1
- SOMMERVILLE R. A., & TAFFLER R. J., (1995) **Banker judgment versus formal forecasting models: the case of country risk assessment**. *Journal of Banking & Finance*, 19, 281-297
- SHARMA, S. (1996). **Applied Multivariate Techniques**. New York: John Wiley & Sons; ISBN 0-471-31064-6
- SCHREINER, M. (1999) **A Scoring Model of the risk of Costly Arrears at a Micro financeLender in Bolivia**. *Washington University in St. Louis.: U.S.A.*
- SCHREINER, M. (2000) **Credit Scoring for Microfinance: Can It Work?**

*Washington University in St. Louis.: U.S.A.*

- THANH D. T. H. & KLEIMEIER S., (2006) **Credit Scoring for Vietnam's Retail Banking Market: Implementation and Implications for Transactional versus Relationship Lending.** *Limburg Institute of Financial Economics (LIFE). Faculty of Economics and Business Administration (FdEWB)*
- THOMAS L. C., (2000) **A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers.** *International Journal of Forecasting* ,16, 149–172
- THOMAS, L. C., EDELMAN, D. B., & CROOK, J. N. (2002). **Credit Scoring and its Applications.** *Siam: Monographs on Mathematical Modeling and Computation.*
- VIGANÓ, L. (1993) Savings and Development: **A Credit Scoring Model for development banks: an African Case Study**, 441-482.
- WOOLDRIDGE, Jeffrey. M. (2002) **Introductory Econometrics: A modern approach.** Publisher: South-Western.
- YAP, B. W. & ONG, S. H., & HUSAIN, N. H. M. **Using data mining to improve assessment of credit worthiness via credit scoring models.** *Expert Systems with Applications Journal* 38, 2011. (13274-13283).