
*Instituto Superior de Economia e Gestão
Universidade Técnica de Lisboa*

Mestrado em Gestão e Avaliação Imobiliária

Determinantes do “Default” no Crédito Habitação Hipotecário

Nuno Miguel da Silva Pinhão Dâmaso Fazenda

Júri:

Presidente: Doutor João Luís Duque, professor catedrático do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;

*Vogais: - Doutor Joaquim Montezuma de Carvalho, professor auxiliar convidado do Instituto Superior de Economia e Gestão da Universidade Técnica de Lisboa;
- Mestre Amaro Naves Laia, especialista na área de Gestão Imobiliária.*

*Lisboa
Setembro de 2008*

*Instituto Superior de Economia e Gestão
Universidade Técnica de Lisboa*

Determinantes do “Default” no Crédito Habitação Hipotecário

Dissertação para a obtenção do grau de mestre
em Gestão e Avaliação Imobiliária

Nuno Miguel da Silva Pinhão Dâmaso Fazenda

Orientador: Dr. Joaquim Montezuma

*Lisboa
Setembro de 2008*

Agradecimentos

Foram várias as pessoas que colaboraram no desenvolvimento deste estudo, às quais não poderia deixar de manifestar o meu profundo e sincero agradecimento, em especial:

Ao meu orientador Dr. Joaquim Montezuma, pela disponibilidade que manifestou na leitura dos textos, acompanhamento e condução do trabalho.

Ao Prof. Dr. João Andrade e Silva, pela pronta disponibilidade e colaboração no tratamento estatístico dos dados.

Ao Dr. Rui Barata pela disponibilidade e ajuda na recolha de dados, que serviram de base ao desenvolvimento da tese.

À minha namorada pelo apoio, incentivo e ajuda prestada durante todo o desenvolvimento da tese.

Aos meus amigos e colegas, pelo apoio e solidariedade manifestada.

A todos aqueles que directa ou indirectamente contribuíram para a boa conclusão desta tese,

o meu MUITO OBRIGADO!

Resumo

Considerando a crescente evolução de crédito, verificada nos últimos anos, nomeadamente no crédito à habitação, acompanhada recentemente por um aumento do incumprimento, julgámos que seria importante a realização de um trabalho, que tentasse compreender os factores, mais determinantes, ou que mais se destacariam, nas situações de incumprimento.

Com o novo Acordo de Basileia a determinação destes factores tem um carácter cada vez mais importante para as instituições de crédito, nomeadamente para as que utilizam métodos de *ratings* internos na análise de risco de crédito.

Neste sentido, realizámos junto de uma entidade bancária nacional, uma recolha de dados, relativos às operações de crédito à habitação efectuadas nos últimos quatro anos, de forma a obter uma amostra com o conjunto das variáveis, que melhor retratassem as respectivas operações.

Na análise dos dados, utilizámos o Método de Regressão Logística (*Logit*), o qual, a partir de um conjunto de variáveis independentes, estima a probabilidade de ocorrer um certo evento, neste caso existência de incumprimento.

A regressão logística minimiza o número de variáveis independentes, determinando as que devem ser incluídas no modelo, de forma a obter um modelo final, mais ajustado e que preveja adequadamente a variável dependente. Aplica-se quando a variável dependente é dicotómica, assumindo valor 1 (existência de incumprimento) ou valor 0 (inexistência de incumprimento).

Com a aplicação do *Logit* no nosso estudo, verificámos que o modelo final, apenas continha quatro das nove variáveis que compunham a amostra. Assim, com base nos dados que caracterizavam a nossa amostra, foram estas quatro variáveis - Taxa de Juro, Prazo do Empréstimo, Taxa de Esforço e Loan-to-Value - as que melhor explicaram a existência de incumprimento nas operações de crédito em análise.

Acreditamos que a existência de uma boa base de dados, agrupando um conjunto de variáveis que retrate fidedignamente as operações de crédito e o ambiente sócio-económico do cliente, poderá, a aplicação do Modelo de Regressão Logística, constituir uma mais valia na determinação dos factores de *default* no crédito habitação hipotecário.

Abstract

Given the growing rise in credit over the last few years, namely in housing credit, recently accompanied by an increase in missed payments, we thought it relevant to produce a study that would attempt to understand the most determining factors (or those which would stand out the most) leading to failure to comply with payments.

In light of the new Basil Agreement, the pin-pointing of such factors is increasingly important for credit providers, especially for those who use internal rating methods when carrying out credit risk assessment.

To that purpose, we gathered data from a national bank on housing credit operations carried out in the last four years, in order to obtain a sample which would include the sets of variables that might best typify the different operations.

We did so by using the Logistical Method of Regression (*Logit*), which uses sets of independent variables to estimate the probability of a given event occurring – in this case, failure to comply with payments. Logistical regression minimizes the number of independent variables by identifying those that should be included in the model, so as to obtain a better adjusted final model, which might adequately predict the dependent variable. It is applicable when the dependent variable is dichotomous, taking the value of 1 (non-compliance present) or of 0 (non-compliance not present).

The use of *Logit* in our study resulted in a final model that contained only four of the nine variables making up the sample. Therefore, based on the data that made up our sample, it was these four variables – (Interest Rate, Duration of the Loan, Effort Rate and Loan-to-Value) – that best explained the existence of non-compliance in the credit operations under analysis.

We believe that the application of the Logistical Regression Model to a good database, which brings together a set of variables that provide an accurate picture of the credit operations and of the socio-economic environment of the client, brings added value in determining the factors leading to defaulted payments in housing credit mortgages.

Índice Geral

AGRADECIMENTOS	2
RESUMO	3
ABSTRACT	4
ÍNDICE GERAL	5
ÍNDICE DE FIGURAS	6
ÍNDICE DE QUADROS	7
ÍNDICE DE GRÁFICOS	8
CAPÍTULO 1	9
1. INTRODUÇÃO	9
2. REVISÃO BIBLIOGRÁFICA	11
2.1 <i>Resenha da Evolução nos últimos Anos</i>	11
2.1.1 Evolução do crédito em Portugal a partir da década 90.....	11
2.1.2 Impacto do agravamento das condições de crédito.....	14
2.2 <i>Factores Determinantes do Incumprimento</i>	16
2.3 <i>Formas de Avaliar o Risco</i>	23
CAPÍTULO 2	24
3. METODOLOGIA.....	24
3.1 <i>Segmentação da População e Obtenção da Amostra</i>	25
3.2 <i>Análise Descritiva</i>	31
3.3 <i>Métodos utilizados na Análise de Default</i>	32
3.1.1 Modelo de Regressão Logística - Logit.....	34
CAPÍTULO 3	42
4. RESULTADOS	42
4.1 <i>Análise Descritiva da Amostra</i>	42
4.1.1 Valor da Prestação.....	44
4.1.2 Taxa de Juro.....	46
4.1.3 Prazo do Empréstimo.....	48
4.1.4 Idade do Cliente.....	51
4.1.5 Taxa de Esforço.....	54
4.1.6 Loan - to - Value	56
4.1.7 Valor do Imóvel.....	58
4.1.8 Rendimento Anual do Agregado Familiar.....	61
4.1.9 Indexante.....	64
4.2 <i>Modelo de Regressão Logística - Logit</i>	66
DISCUSSÃO	78
BIBLIOGRAFIA	82

Índice de Figuras

FIGURA 1 - SEGMENTAÇÃO DA POPULAÇÃO DE CRÉDITO À HABITAÇÃO	26
FIGURA 2 - SISTEMATIZAÇÃO DO PROCESSO DE RECOLHA DE DADOS.....	29
FIGURA 3 - MÉTODOS DE ANÁLISE DESCRITIVA UTILIZADOS	31
FIGURA 4 - CURVA LOGÍSTICA	35

Índice de Quadros

QUADRO 1 - CLASSIFICAÇÃO DO CRÉDITO DA AMOSTRA - OPERAÇÕES MOROSAS E NÃO MOROSAS	42
QUADRO 2 - CLASSIFICAÇÃO DO CRÉDITO DA AMOSTRA - DIAS EM ATRASO	43
QUADRO 3 - ESTATÍSTICA DA VARIÁVEL "VALOR DA PRESTAÇÃO"	44
QUADRO 4 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "VALOR DA PRESTAÇÃO"	46
QUADRO 5 - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "TAXA DE JURO"	47
QUADRO 6 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "TAXA DE JURO"	48
QUADRO 7 - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "PRAZO DE EMPRÉSTIMO"	49
QUADRO 8 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "PRAZO DO EMPRÉSTIMO"	50
QUADRO 9 - ESTATÍSTICA DA VARIÁVEL "IDADE DO CLIENTE"	52
QUADRO 10 - - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "IDADE DO CLIENTE"	53
QUADRO 11 - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "TAXA DE ESFORÇO"	54
QUADRO 12 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "TAXA DE ESFORÇO"	55
QUADRO 13 - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "LOAN TO VALUE"	56
QUADRO 14 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "LOAN TO VALUE"	58
QUADRO 15 - - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "VALOR DO IMÓVEL"	59
QUADRO 16 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "VALOR DO IMÓVEL"	60
QUADRO 17 - ESTATÍSTICA DESCRITIVA DA VARIÁVEL "RENDIMENTO ANUAL DO AGREGADO FAMILIAR"	62
QUADRO 18 - QUANTIFICAÇÃO DAS OPERAÇÕES MOROSAS (TERMOS ABSOLUTOS E RELATIVOS) PARA A VARIÁVEL "RENDIMENTO ANUAL"	63
QUADRO 19 - CLASSIFICAÇÃO DO INDEXANTE - FREQUÊNCIA DE INCUMPRIMENTO	65
QUADRO 20 - CASE PROCESSING SUMMARY	67
QUADRO 21 - - CLASSIFICACION TABLE A,B	67
QUADRO 22 - OMNIBUS TESTS OF MODEL COEFFICIENTES	68
QUADRO 23 - MODEL SUMMARY	69
QUADRO 24 - HOSMER AND LEMESHOW TEST	70
QUADRO 25 - CLASSIFICACION TABLE A	71
QUADRO 26 - AMOSTRA DE VALIDAÇÃO (CLASSIFICACION TABLE A)	74
QUADRO 27 - AMOSTRA DE TREINO (CLASSIFICACION TABLE A)	74
QUADRO 28 - VARIABLES IN THE EQUATION	75
QUADRO 29 - INTERVALOS DE PROBABILIDADE DE OCORRER INCUMPRIMENTO	77

Índice de Gráficos

GRÁFICO 1 - DIVERSIFICAÇÃO DO CRÉDITO DA AMOSTRA - ESTADO DAS OPERAÇÕES	43
GRÁFICO 2 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA VARIÁVEL "VALOR DA PRESTAÇÃO"	44
GRÁFICO 3 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "VALOR DE PRESTAÇÃO"	45
GRÁFICO 4 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA A VARIÁVEL "TAXA DE JURO"	46
GRÁFICO 5 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA INTERVALOS DE "TAXA DE JURO"	47
GRÁFICO 6 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA VARIÁVEL "PRAZO DE EMPRÉSTIMO"	48
GRÁFICO 7 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "PRAZO DO EMPRÉSTIMO"	50
GRÁFICO 8 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA VARIÁVEL "IDADE DO CLIENTE"	51
GRÁFICO 9 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "IDADE DO CLIENTE"	52
GRÁFICO 10 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA A VARIÁVEL "TAXA DE ESFORÇO"	54
GRÁFICO 11 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "TAXA DE ESFORÇO"	55
GRÁFICO 12 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA A VARIÁVEL "LOAN TO VALUE"	56
GRÁFICO 13 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "LOAN TO VALUE"	57
GRÁFICO 14 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA A VARIÁVEL "VALOR DO IMÓVEL"	58
GRÁFICO 15 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "VALOR DE AVALIAÇÃO"	60
GRÁFICO 16 - REPRESENTAÇÃO GRÁFICA DOS DADOS DA AMOSTRA PARA VARIÁVEL "RENDIMENTO ANUAL"	61
GRÁFICO 17 - FREQUÊNCIA DE OPERAÇÕES MOROSAS PARA DIFERENTES INTERVALOS DE "RENDIMENTO ANUAL DO AGREGADO FAMILIAR"	62
GRÁFICO 18 - CLASSIFICAÇÃO DO INDEXANTE POR PRAZO DE INDEXAÇÃO	64
GRÁFICO 19 - FREQUÊNCIA DE INCUMPRIMENTO CONSOANTE O INDEXANTE	65

Capítulo 1

1. Introdução

De uma forma generalizada e global, a conjuntura verificada nos últimos 20 anos permitiu um despoletar de níveis substanciais de concessão de crédito por parte das instituições financeiras, as quais, por força de um mercado concorrencial e cada vez mais competitivo, têm concedido crédito de forma massiva e sem uma correcta análise à capacidade financeira dos mutuários para levar a efeito o serviço da dívida, originando uma situação de incumprimento (ou *default* - estrangeirismo adoptado na cena económica para traduzir incumprimento).

O Incumprimento, pode ser definido como o não pagamento dos compromissos financeiros por parte de um mutuário e normalmente advém de situações em que existe endividamento (endividamento traduz o saldo devedor de uma determinada entidade, após o respectivo recurso ao crédito), independentemente da forma em que este ocorre.

Geralmente as instituições financeiras consideram que há incumprimento de crédito ao fim de três prestações em atraso (no caso de prestações mensais será ao fim de 90 dias) e incumprimento definitivo quando se esgotam as possibilidades de renegociação e se iniciam procedimentos judiciais de cobrança coerciva.

Marques et al (2003) definem crédito como a “disponibilização imediata de rendimento que não se possui, permitindo antecipar a fruição de determinados bens, mas implica igualmente uma penhora do rendimento futuro, impondo aos devedores um sacrifício financeiro por períodos de tempo mais ou menos longos”.

Nesta perspectiva pode-se considerar que o crédito é um factor importante na melhoria das condições de vida das famílias, na medida em que lhes permite antecipar a acessibilidade a

determinados bens, mas simultaneamente, obriga-as a um esforço acrescido de gestão do orçamento disponível e a uma poupança forçada, para não correrem o risco de entrar em situações de incumprimento.

De uma maneira geral o crédito aos consumidores compreende todo o empréstimo a particulares que não tem por finalidade uma actividade económica e profissional. Inclui crédito destinado à aquisição de bens imobiliários (crédito à habitação) e crédito destinado à aquisição de outros bens e serviços (crédito ao consumo). A problemática do endividamento das famílias abrange essencialmente estas duas vertentes (Marques e Frade, 2003).

O presente trabalho vai focar principalmente o crédito destinado à aquisição de bens imobiliários (crédito a habitação).

São muitos os factores que poderão estar na origem do incumprimento no crédito a habitação, podendo ser factores inerente à própria operação de crédito (estes mais fáceis de detectar), mas também exteriores às operações, nomeadamente de carácter pessoal/familiar, ou social.

A título de exemplo podemos referir que a literatura, a partir de alguns estudos já realizados tem apontado como factores inerentes à própria Operação, a taxa de juro, prazo do empréstimo, Loan-to-value, prazo do indexante; inerentes à Pessoa/Família a idade, estado civil, sexo, tamanho do agregado, estado de saúde; de carácter Social, o rendimento, situação perante o emprego (vinculo laboral, tempo e duração de contrato). Contudo, alguns destes factores podem estar interligados e quando verificados em conjunto, explicarem melhor a razão da existência de incumprimento.

2.Revisão Bibliográfica

2.1 Resenha da Evolução nos últimos Anos

2.1.1 Evolução do crédito em Portugal a partir da década 90

Em Portugal, a expansão do crédito aos consumidores deu-se na década dos anos noventa, acompanhando a liberalização dos mercados financeiros e o aumento do rendimento das famílias. A inexistência de um mercado de arrendamento e a modernização da oferta comercial contribuíram também para a expansão deste fenómeno. (Marques *et al.*, 2000).

Assistiu-se então a um aumento do património dos particulares, comparativamente com o seu rendimento disponível, tendência crescente e que se manteria durante toda essa década. Assim o crescimento da riqueza das famílias, nomeadamente patrimonial, teve como contrapartida um aumento muito significativo do endividamento, sobretudo no que se refere a empréstimos de longo prazo, destinados à aquisição de habitação (Cardoso e Cunha 2005).

O acréscimo de procura de habitação com recurso ao crédito foi especialmente marcado pela redução acentuada das taxas de juro, pela maior facilidade de acesso ao crédito por parte dos particulares e pela liberalização do mercado financeiro, donde resultou maior concorrência entre instituições bancárias, provocando assim diminuição de *spreads*, o que era convidativo à aquisição do crédito.

Nesta década, verificou-se ainda uma redução da inflação e um crescimento continuado do Produto Interno Bruto (PIB), o que ajudou à tendência de um maior endividamento das famílias, quer no crédito à habitação quer no crédito ao consumo, agravando a taxa de esforço das famílias para o cumprimento dos seus compromissos de crédito (CM Lisboa 2005).

Efectivamente, desde meados dos anos 90, que o rácio de endividamento total das famílias portuguesas em relação ao PIB tem registado elevadas taxas de crescimento (de 56,3% em 2000 aumentou para 68,00% em 2005, donde se exclui o crédito titularizado), bem como o endividamento hipotecário (de 41,5% em 2000 para 54,0% em 2005 - exclui o crédito titularizado), constituindo a principal razão do endividamento privado em Portugal (Genworth Financial, 2006).

A compra ou construção de casa própria para residência permanente constitui para grande parte das famílias o maior esforço financeiro que toda a sua vida. Diversos factores poderão ter estado na origem do aumento de importância da habitação própria: restrições da oferta no mercado de arrendamento privado, exiguidade do sector da habitação social e um claro direccionamento da política de habitação via crédito bonificado (para incentivar a aquisição de habitação entrariam em vigor incentivos aos jovens através da bonificação de juros). (CM Lisboa 2005).

Ainda durante a década de 90, principalmente na segunda metade, verificou-se simultaneamente três fenómenos, um aumento dos valores do stock de habitação, um aumento dos seus preços relativos, e uma descida das taxas de juro.

A conjugação destes três factores, incrementou o incentivo ao investimento em habitação, onde para além de existir aumento da oferta, verificar-se por sua vez, um aumento dos preços relativos da habitação também reflecte o crescimento em alta da procura de habitação.

A taxa de crescimento do stock de habitação foi sempre superior à da população, reflectindo dois tipos de factores. Por um lado, a redução na dimensão média dos agregados familiares e o consequente aumento do seu número: de acordo com dados censitários do Instituto Nacional de Estatística (INE), o número médio de pessoas por família passou de 3.1 em 1991 para 2.8 em 2001, registando-se um particular crescimento das famílias unipessoais entre 1991 e 2001, passando a representar 17% do total, quando em 1991 representavam 14%. Por outro lado, a expansão do parque habitacional reflecte também o aumento do número de alojamentos por família (1.4 em 2001 e 1.3 em 1991) associado à crescente importância dos alojamentos de uso sazonal, que em 2001 correspondiam a 18% dos alojamentos face a 16% em 1991 (Cardoso e Cunha 2005).

Terminada a década de 90, verificou-se a partir de 2000, um abrandamento no endividamento dos particulares, nomeadamente a partir de 2002, ano em que o crédito bonificado deixou de existir.

Este abrandamento no endividamento ocorreu durante um curto período de tempo, no qual o mercado se ajustou e respondeu a este abrandamento.

Como resposta a este abrandamento, uma das formas de compensar os potenciais beneficiários deste regime de crédito, e continuar a promover a aquisição de habitação, foi aumentar o prazo máximo dos empréstimos de crédito à habitação, que até à altura era de 30 anos.

Por outro lado, a Euribor a 6 meses, indexante de referência dos empréstimos à habitação em Portugal, iniciou em 2005 uma trajetória ascendente, contribuindo para um acréscimo dos encargos das famílias portuguesas junto da Banca. Acrescentando ainda os reduzidos níveis de confiança dos consumidores, o aumento da taxa de desemprego, entre outros factores, verificou-se assim, um consumo privado, com uma taxa de variação superior à do rendimento disponível. Foi notória, nesta data, a nova diminuição da taxa de poupança dos particulares e o aumento do endividamento (Genworth Financial 2006).

Para justificar este novo aumento do endividamento há evidências que sugerem ter existido alguma adaptação do lado da oferta de crédito à capacidade corrente dos clientes para assegurar o serviço da dívida. A introdução de produtos no mercado com características que permitem o diferimento do grau de esforço associado terá certamente contribuído para a evolução observada no segmento dos particulares em Portugal (Banco Portugal 2005).

De acordo com a informação divulgada pela Direcção Geral do Tesouro, verificou-se em 2005 um aumento do número de contratos de crédito na ordem dos 7,8%, o qual poderá estar associado à prática de rácios *loan-to-value* menos exigentes, emprestando os bancos valores superiores ao valor do imóvel, por forma a facilitarem a aquisição de habitações de preço superior. Este aumento deverá traduzir não apenas a aquisição de nova habitação mas também a realização de novos contratos para substituição de empréstimos obtidos anteriormente em condições menos vantajosas (Banco Portugal 2005).

2.1.2 Impacto do agravamento das condições de crédito

Alguns estudos têm apontado para uma correlação entre a evolução do preço das habitações e o consumo, na medida em que o aumento do preço da casa representa para o proprietário um aumento de riqueza, conferindo-lhe uma capacidade acrescida de acesso ao crédito e, conseqüentemente, capacidade para consumir. Mesmo que o proprietário não aumente o montante em crédito, ao ver o seu activo valorizado, sente-se mais rico, estando assim mais propenso a consumir (Quigley, 1997)

Num estudo do Centro de Estudos Financeiros da Universidade de Frankfurt (2006), analisou-se a correlação entre os preços das casas e a evolução do consumo nalguns países. Nos Estados Unidos essa correlação é bastante elevada devido ao uso generalizado de produtos de refinanciamento hipotecário que permitem aos proprietários das casas aumentarem facilmente o seu financiamento à medida que estas vão valorizando. Em Espanha, a correlação entre o consumo e os preços das casas é ainda maior (Calza et al. 2006).

Se o preço das casas neste país continuar a desacelerar, iremos assistir a um abrandamento do consumo naquele país. Esta é uma situação delicada, pois pode ter graves repercussões em Portugal, dada a interdependência que o nosso país possui com a economia de vizinha Espanha.

O crédito malparado (crédito que os bancos ainda não conseguiram recuperar) atingiu o valor mais alto de sempre no final de Agosto de 2007. Dados do Boletim Estatístico de Outubro de 2007 do Banco de Portugal, apontam um total de 97,9 mil milhões de euros de financiamento concedido pelos bancos aos particulares para compra de casa, tendo ficado por pagar mais de 1,2 mil milhões de euros, traduzindo um aumento de 5% de crédito malparado, quando comparado com os primeiros oito meses de 2006.

Veja-se o caso do *subprime*, que está na base da actual crise do mercado imobiliário americano e que teve origem em operações de crédito de risco elevado, feitas com clientes de baixos rendimentos e com uma situação económica instável.

Em 2003, a economia americana apresentava uma taxa de juro real de 1%, a criação de emprego e o investimento empresarial estavam em níveis muito baixos, levando os bancos a ser menos exigentes na concessão de crédito, realizando empréstimos a 100% do valor das avaliações e operações com taxa de esforço dos clientes elevada.

O problema surgiu quando o Reserva Federal Norte-Americana começou a subir de novo os juros, aumentando de 1% para 5% em dois anos. Os preços das casas entraram em queda, as prestações das casas subiram e grande parte destas famílias, cujos rendimentos não acompanharam estas subidas, ficou sem capacidade para pagar os empréstimos.

No decurso deste período existiu um grande aumento dos preços das habitações, a que se seguiu uma inevitável fase de correcção. Para além dos particulares, também muitos investidores entraram em incumprimento, uma vez que o aumento dos custos de financiamento (por subida das taxas) diminuiu o retorno dos investimentos especulativos que tinham realizado.

Se efectuarmos uma comparação com o ocorrido na zona euro nos últimos anos, verificamos algumas semelhanças, nomeadamente, na zona euro a valorização das casas foi superior à dos Estados Unidos, tendo nos últimos sete anos os preços das casas subido 89% na UE e 79% nos EUA. Além disso, as taxas de juro também aumentaram bastante na economia europeia, apesar de menos do que nos Estados Unidos, tendo a taxa passado de 2,0% para 4,0% no último ano e meio. À primeira vista, estes dados sugerem que estamos perante um forte risco de abrandamento do mercado imobiliário e consequentemente de abrandamento da economia (Dwight et al, 2007).

A nossa economia pode ser afectada por estes desenvolvimentos externos, caso as condições de crédito sejam apertadas no seguimento da crise, dada a sua forte dependência do exterior.

Acresce ainda que o crédito à habitação em Portugal é muitas vezes concedido sem uma correcta avaliação da capacidade financeira das famílias, ainda que haja uma garantia do pagamento através da hipoteca da habitação, o que por enquanto ainda constitui uma garantia real do seu pagamento.

Os problemas poderão começar a surgir quando, à imagem do que se verifica nos Estados Unidos, os preços das habitações começarem a baixar e o valor das hipotecas deixarem de

cobrir o valor dos empréstimos. Numa situação destas, as próprias instituições de crédito terão dificuldades em accionar as garantias das habitações em situações de incumprimento. Não deixa de ser preocupante, dados os elevados níveis de endividamentos das nossas famílias, um dos mais elevados da Europa, e pelo crescente do crédito de cobrança duvidosa no total dos empréstimos bancários (Banco de Portugal).

O aumento das taxas de juro e dos *spreads* poderá ter consequências gravíssimas em toda a economia da zona euro. Os credores irão enfrentar maiores dificuldades financeiras com o aumento das suas prestações, principalmente países com empréstimos maioritariamente indexados a taxa variável, sendo muito provável um aumento do número de incumprimentos.

Será ainda provável que o recurso ao crédito venha a diminuir, quer por via da contenção da oferta, dado que os bancos estão mais relutantes em ceder fundos, quer por via da procura.

Embora as actuais restrições à concessão de crédito não sejam suficientes para pôr em causa o crescimento económico, o crédito pode tornar-se mais caro e difícil de obter, o que poderá ter um impacto negativo na confiança dos consumidores originando quebras do consumo, bem como na expansão do negócio das empresas.

2.2 Factores Determinantes do Incumprimento

Pela análise de diversos estudos sobre a matéria, constata-se que o incumprimento pode atingir com maior incidência determinadas franjas da sociedade. Podem estar na origem do incumprimento factores pessoais, económicos, sociais ou intrínsecos às próprias operações de crédito. A idade, sexo, estado civil, taxas de juro, rácio valor do empréstimo/valor do imóvel (*loan-to-value*), período do empréstimo, entre outros, podem, em maior ou menor número, indiciar potenciais situações de risco de incumprimento.

O Observatório do Endividamento dos Consumidores (OEC) tem realizado um conjunto de estudos sobre a evolução do incumprimento no crédito a habitação.

Com base nesses estudos, verificou-se que o incumprimento e o rácio de incumprimento (rácio entre o montante de capital vencido e o total dos saldos devedores), relativos ao crédito a habitação, têm mantido uma trajectória ascendente. Refira-se algumas conclusões divulgadas no seu Relatório de Actividades de 2002:

“O incumprimento é particularmente importante nos distritos de Setúbal, Lisboa e Porto, representando nestes três distritos cerca de 71,4% do total de situações de incumprimento no país.

Entre os três principais regimes de crédito à habitação, o incumprimento é relativamente mais frequente no caso dos regimes bonificado e jovem bonificado do que no regime geral.

O incumprimento é mais significativo no caso de mutuários separados judicialmente, divorciados e solteiros. Relativamente à faixa etária, o incumprimento é também relativamente mais observável, embora as diferenças sejam pouco sensíveis, no caso de mutuários com idades entre os 36 e os 45 anos.

Refira-se ainda que os contratos em regime de prestações crescentes apresentavam uma taxa de incumprimento superior à verificada nos contratos em regime de prestações constantes”.

Um desses trabalhos, incidiu sobre a análise ao nível de sobreendividamento dos consumidores, assente em 46 inquéritos recebidos da DECO, realizados na região de Coimbra, entre Janeiro de 2001 e Agosto de 2002, no qual se traçou o perfil do sobreendividado: A maioria dos sobreendividados que pediram auxílio à DECO de Coimbra eram sobretudo do sexo masculino, situavam-se na faixa etária entre os 35 e os 55 anos, eram casados e possuíam o ensino secundário (45,7%), sendo também significativo o número de pessoas com formação superior (21,7%). Predominavam também os trabalhadores da Administração pública, seguidos de comerciantes e vendedores e dos reformados. O rendimento do agregado familiar dos requerentes situava-se entre 750 a 1.250 euros/mês, seguindo-se a classe com rendimentos de 1.250 a 2.000 euros. A classe de rendimentos mais baixa, até 75 euros, surge apenas em terceiro lugar.

O estudo aponta como principais causas do sobreendividamento o desemprego, a escassez de rendimentos e problemas de saúde.

Concluiu-se também que o aumento das prestações dos créditos tem um peso muito pouco significativo no sobreendividamento, o que permite supor que, em geral, as famílias possuem alguma margem financeira para responder a pequenas oscilações no valor da taxa de juro.

Resumidamente, os dados apresentados pelo OEC, mostram que foram uma combinação de motivos que conduziram as pessoas ao seu problema financeiro: desemprego, insuficiência de rendimento, má gestão do orçamento familiar disponível, problemas de saúde, alteração do agregado familiar e agravamento do custo do crédito.

Informações recolhidas junto do Gabinete de Apoio ao Sobreendividamento da DECO, mostram que as pessoas singulares que recorrem a esse serviço são normalmente pessoas casadas, com idades compreendidas entre os 25 e os 55 anos, um nível de instrução e de rendimento médios, vários créditos (multiendividamento), a quem a perda do emprego ou os problemas de saúde, aliados à insuficiência de rendimento e à má gestão do orçamento familiar, provocaram a ruptura financeira (Marques e Frade).

Num outro estudo realizado na Faculdade de Economia da Universidade de Coimbra (FEUC) - Marques et. al, constatou-se que o recurso ao crédito para consumo é maior por parte dos homens (21,9 %) do que por parte das mulheres (16,1 %), sendo por classe etária mais frequente no grupo dos 35 aos 44 anos (30,1 %), seguindo-se o grupo dos 45 aos 54 (24,8 %). Verificou-se ainda que os indivíduos das localidades mais pequenas recorrem menos ao crédito que os indivíduos dos grandes centros urbanos, o mesmo acontecendo com os pertencentes às classes baixas (de 15,7 % nos segmentos mais baixos, passa-se para 25,5 % nas classes altas). No entanto, são os dos meios desfavorecidos aqueles que têm mais dificuldade em pagar as dívidas. Possíveis alterações da situação familiar (divórcio), do rendimento (desemprego) ou do património são algumas das causas apontadas para o incumprimento.

À semelhança dos anteriores estudos apresentados, este conclui que a taxa de incumprimento está fortemente associada à natureza dos empréstimos ou às características socioeconómicas dos mutuários, sendo maior nos regimes de crédito bonificado, no regime de amortização em prestações crescentes, nos escalões de rendimentos mais baixos e em situação de divórcio e separação judicial. Aponta igualmente que as situações de incumprimento estão concentradas nos distritos de Lisboa, Porto e Setúbal.

Tendo por base de estudo o tema *default*, Yongheng et al, (Março de 2000), confirmam também que determinados eventos difíceis na vida do mutuário, tais como o desemprego (ainda que se perspetive nova colocação a curto prazo), e o divórcio, poderão explicar um comportamento de *default*.

No que diz respeito ao factor idade dos mutuários o Banco de Portugal tem alertado para o facto do aumento do acesso ao crédito por parte dos indivíduos mais jovens ter vindo a aumentar e a ser muito expressivo, recomendando prudência na sua concessão, uma vez que este segmento da população possui rendimentos mais baixos e uma maior propensão para transitar para situações de desemprego (Genworth Financial).

Efectivamente, na segunda metade dos anos noventa, foram os "mais jovens" os que mais contribuiram para o aumento do endividamento agregado. Este grupo é o que "apresenta as taxas de esforço médias mais elevadas" e o "mais sub-representado nos inqueritos" (Banco de Portugal).

Num estudo realizado no Brasil por Zerbini e Rocha (2004) concluiu-se que na análise ao comportamento de incumprimento, as características pessoais não são tão importantes quanto as características dos créditos concedidos. O valor das prestações é o principal determinante na probabilidade de incumprimento, seguido pelo número de prestações, o que mostra que o aumento dos períodos de vencimento só terá contribuição positiva no controle de casos de incumprimento quando os juros cobrados e, portanto, o valor das prestações, for significativamente reduzido. Defendem estes autores que a idade, sexo, casa própria, o número de anos na actual residência e o número de anos no actual emprego não são estatisticamente significativas. No entanto é possível traçar o perfil de bom pagador: casado, trabalha no sector público.

Complementando, de certa forma a ideia atrás referida, de que o incumprimento advém das condições inerentes ao empréstimo, Fuinhas (2004) refere que com taxas de juro nominais e reais mais baixas, o fardo da dívida reparte-se mais uniformemente pela duração do empréstimo, favorecendo-se neste sentido as condições de cumprimento.

Salienta ainda, que em períodos de inflação elevada o esforço da dívida, em termos reais, é mais elevado no início do empréstimo e reduz-se rapidamente à medida que o tempo decorre (Fuinhas 2004).

Outro aspecto a ter em conta para a possibilidade de ocorrência de incumprimento advém do facto de em algumas ocasiões as taxas de juro aplicadas pelos mutuantes serem uma função crescente do grau de endividamento dos mutuários. Numa situação limite, os intermediários financeiros podem não querer conceder crédito, mesmo que os mutuários se manifestem predispostos a pagar taxas de juro mais elevadas (Fuinhas 2004). Neste caso, a concederem, estariam a realizá-lo por vezes em condições acima das capacidades dos mutuários, o que agravaria o risco de incumprimento.

Embora referindo-se a operações de crédito hipotecário em empresas, outros autores afirmam que a evolução da actividade macroeconómica desempenha também um papel bastante relevante na avaliação das probabilidades de incumprimento ao longo do tempo, sugerindo a existência de uma clara relação entre o risco de crédito e a evolução da situação macroeconómica (Pederzoli e Torricelli (2005), Jiménez e Saurina (2006) ou Bonfim (2007)).

Bonfim conclui exactamente que, embora as probabilidades de incumprimento sejam determinadas essencialmente pelas características específicas de cada empresa, existe igualmente uma relação entre a evolução da actividade macroeconómica e as taxas de incumprimento (Bonfim, 2006).

À semelhança do apresentado nos casos de financiamento destinado à actividade empresarial, pensamos que as condições económicas podem igualmente influenciar e levar a situações de incumprimento no crédito hipotecário para habitação.

Ronel Elul (2006), confirma que a evolução negativa do mercado imobiliário habitacional, ou seja, a diminuição dos preços dos imóveis, constitui um factor de incumprimento dos empréstimos hipotecários.

Alguns economistas desenvolveram modelos de opções que permitem mesmo quantificar o impacto que a diminuição dos preços tem no incumprimento do crédito hipotecário (Elul 2006).

Kau *et al* (Maio de 1994) acrescentam que os preços mais voláteis das casas estão associados a uma maior incidência de probabilidade de incumprimento.

Para além dos factores inerentes às próprias operações de crédito, atrás referidos, há alguns autores que defendem que o rácio entre o valor do crédito e o valor do imóvel - Loan-to-value (LTV) – é um bom determinante da probabilidade de incumprimento. Outros autores, por seu lado discordam desta teoria, apresentando no entanto outras vantagens para a análise deste rácio.

Segundo a revista Genworth Financial a probabilidade de incumprimento é maior em empréstimos cujo rácio LTV é mais elevado, comparativamente a situações de baixo LTV.

Neste sentido as estimativas da Genworth Financial com base nos dados da agência de rating Fitch Ratings apontam o LTV como um indicador fiável da probabilidade de incumprimento.

A taxa de incumprimento de empréstimos com LTVs entre 85% e 90% é, em média, 1,26 vezes mais alta que nos empréstimos com LTVs entre 80% e 85%, sendo, em LTVs de 90% a 95%, 1,58 vezes superior.

A experiência dos EUA vem igualmente confirmar a importância do LTV como indicador do risco creditício, mesmo em períodos de resultados macroeconómicos favoráveis. Por exemplo, apesar das taxas de desemprego mais baixas da história daquele país e de um forte aumento no preço da habitação, o incumprimento para LTVs entre 80% e 90% continuam a ser 3 e 4 vezes maiores que as de empréstimos com LTVs de 75% a 80%.

Contrapondo de certa forma o atrás referido, Archer *et al* (Junho de 1999) defendem não haver uma relação empírica observável entre o LTV e a taxa de incumprimento. Sugerem que a capacidade de cumprir com o serviço da dívida é mais importante para explicar um comportamento de incumprimento que o LTV. Sugerem ainda que existe uma relação bem mais próxima entre as características do empréstimo e a taxa de incumprimento.

À semelhança destes últimos, Sheridan *et al*, consideram também que a relação entre o LTV e o rácio de incumprimento é relativamente fraca.

Estes autores vêm no entanto referir a existência de outras relações entre o LTV e as condições dos empréstimos. Segundo eles, o rácio médio de LTV por credor possui uma forte relação com os spreads praticados, donde se constata que em empréstimos com LTV elevados, os mutuantes normalmente requerem spreads mais altos, independente do nível de risco da operação (Sheridan et al. 2005).

Segundo alguns autores o problema do endividamento deve ser analisado sobe a perspectiva de outro indicador de grande importância: Taxa de Esforço

Esta tem sido referida nas publicações do Banco de Portugal como o quociente entre a estimativa dos encargos com dívidas (juros e amortizações) suportados pelas famílias e a estimativa do rendimento disponível do agregado.

A determinação da taxa de esforço permite aferir qual o peso da prestação do empréstimo no rendimento mensal disponível das famílias

Segundo Farinha (2004), durante os anos 90 verificou-se a existência de uma diminuição da taxa de esforço das famílias, sendo mais um factor que impulsionou ao recurso ao crédito. Acontece que esta diminuição da taxa de esforço verificada não esteve associada a aumentos de rendimentos do agregado, mas sim à diminuição dos encargos com dívidas, pelo facto da descida das taxas de juro que ocorreram nessa altura.

Em pouco tempo, com o aumento da taxa de juro, assistiu-se à inversão desta situação, tornando-se o valor dos encargos com dívidas (prestações dos empréstimos) a ser maiores e como tal o rendimento disponível relativamente a estes encargos diminui. O elevado valor da taxa de esforço tem condizido a situações de sobreendividamento, e a surgir situações de incumprimento com as prestações dos empréstimos.

O crescimento da taxa de esforço durante este período foi considerável, estando a 9% em 1990 a 8% e em 1999 a 23,5%, sendo de 28% em 2004.

Segundo dados do Banco de Portugal actualmente encontra-se acima sendo de 35% a 40%.

2.3 Formas de Avaliar o Risco

De uma maneira geral a carteira hipotecária analisa-se com base numa série de operações individuais e independentes, onde o que conta é o perfil do proponente e a garantia apresentada. Periodicamente, analisa-se o risco global assumido pela entidade financeira dentro de um perfil de cliente determinado, numa zona geográfica específica, para um intervalo de LTV e para uma determinada taxa de esforço.

Cada uma das operações individuais, é igualmente analisada e avaliado o seu risco. Para este efeito são utilizados complexos modelos e sistemas especializados na avaliação da qualidade da operação de crédito, denominados modelos de *scoring*.

O novo acordo de Basileia – Basileia II, apresenta algumas metodologias de análise do risco, com diferentes graus de sofisticação, incentivando as instituições de crédito a utilizarem as mais avançadas. Assim, as instituições de crédito poderão optar entre a abordagem standard, a abordagem simplificada do método dos *ratings* internos (*IRB Foundation*) ou a abordagem avançada do método dos *ratings* internos (*IRB Advanced*).

O que diferencia a abordagem standard daquelas de *ratings* internos, advém do facto dos factores que influenciam a determinação do risco de crédito serem determinados na segunda por modelos internos e próprios de cada instituição de crédito.

Assim para a melhor utilização dos vários modelos é necessário a definição dos factores que melhor poderão determinar o risco de crédito de forma a prevenir a ocorrência de futuras situações de incumprimento.

O presente trabalho visa identificar e quantificar os principais factores que poderão estar na génese do incumprimento no crédito hipotecário destinado a habitação em Portugal.

Concretamente, pretende-se estudar o impacto no incumprimento dos factores inerentes ao mutuário, às próprias condições das operações e, principalmente, atestar se o valor do imóvel relativamente ao valor do empréstimo (LTV) tem efectivamente influência na probabilidade de *default*.

Capítulo 2

3. METODOLOGIA

Nesta fase da dissertação será referida a forma de obtenção dos dados de modo a criar uma amostra representativa de toda a população, e a metodologia utilizada para tratamento dos dados dessa amostra.

Utilizou-se para obtenção dos dados dois tipos de pesquisa, uma exploratória e outra quantitativa.

Uma pesquisa exploratória, dado o assunto a ser abordado não ser totalmente conhecido (constitui um assunto que ainda poderá ser bastante aprofundado, onde ainda há aspectos por explorar), estando a sua discussão a dar os primeiros passos em Portugal.

Uma pesquisa quantitativa, dado que as informações são mensuráveis, ou seja, todas as informações podem ser apresentadas sob a forma numérica, para subsequentemente serem classificadas e analisadas.

3.1 Segmentação da População e Obtenção da Amostra

Para a realização deste trabalho, foi efectuada uma recolha e tratamento de dados, de forma que a amostra resultante espelhasse da melhor forma possível toda a população.

O presente estudo foi desenvolvido numa instituição de crédito nacional, tendo a recolha incidido apenas sobre um pequeno conjunto de balcões localizados na região centro do país.

Trata-se de uma zona caracterizada por um misto de rural e urbano, onde os clientes do banco são maioritariamente da zona rural e de faixa etária média - alta.

Os dados analisados, não constituem assim uma amostra da totalidade dos dados do banco a nível nacional, mas caracterizam, as operações realizadas a nível regional, por um conjunto dos seus balcões.

Como ponto de partida tomou-se por base todos os dados existentes e definiu-se um período de análise onde a informação estivesse mais completa, de forma a tornar tão fidedigno e representativo quanto possível o objectivo do presente estudo.

Os dados analisados foram recolhidos à data de 31 Dezembro de 2007 e correspondem ao conjunto de operações de crédito realizadas durante um período de quatro anos, num conjunto de cinco balcões daquela instituição.

Em virtude da existência de algumas lacunas na base de dados da instituição, muita da informação relativa às operações e aos clientes não estava disponível para todas as operações realizadas e como tal não constava na base de dados da instituição, advindo daí a necessidade de realizar inúmeras triagens, de forma a constituir uma população base e, a partir desta, uma amostra mais homogénea.

A primeira triagem consistiu em separar as operações de crédito relacionadas com financiamento à habitação, de todas as restantes operações.

Assim, a população inicial de operações de crédito à habitação, foi composta por todas as operações de crédito pertencentes à base de dados daqueles balcões, que não incluíam populações de crédito pessoal, crédito comercial e crédito ao investimento.

A base de dados à qual nos referimos como população inicial, correspondia aos produtos de crédito com as seguintes designações, ou códigos:

- Credito Habitação Regime Geral;
- Crédito Habitação Bonificado;
- Crédito Habitação Jovem Bonificado;
- Crédito Habitação para Aquisição de Propriedades Rurais;
- Crédito Habitação com Carência de Capital;
- Obras de Beneficiação na Habitação;
- Pagamento de Sinal para Aquisição de Habitação;
- Aquisição de Terreno para Construção;
- Obras de Beneficiação na Habitação;
- Transferência de crédito habitação de outras Instituições;
- Aquisição de Habitação Secundária;
- Aquisição de Habitação/Construção para Venda – Investimento.

Como se encontraram diferentes categorias de operações na população de Crédito à Habitação, houve necessidade de as segmentar por grupos, para melhor analisar cada um dos respectivos grupos e eliminar os de pouca representatividade para o nosso estudo.

Assim, a partir da população total de Crédito à Habitação identificámos os seguintes segmentos:



Figura 1 - Segmentação da população de crédito à habitação

Todo este processo de segmentação foi realizado simultaneamente em operações morosas e não morosas.

Foram consideradas operações morosas, ou seja em incumprimento, todas aquelas em que os mutuários estavam em atraso no pagamento à mais de 90 dias.

Com a segmentação, todas as operações referentes a Crédito a Empresas foram eliminadas do processo de modelização, por não fazerem parte do âmbito do estudo.

Relativamente ao Crédito Bonificado/Crédito Não Bonificado, optou-se por uma separação completa aquando da análise dos dados, uma vez que se verificou a existência duma excessiva taxa de incumprimento no primeiro. Considerámos que a análise dos dados em conjunto poderia distorcer qualquer observação e comprometer os resultados.

Acresce ainda o facto de, pela análise efectuada, ter-se verificado que a carteira de Crédito Não Bonificado não sofreu grandes alterações quanto à idade e rendimento médio dos clientes, desde o fim do regime de bonificação em Outubro de 2002, pelo que podemos concluir que a exclusão das operações de crédito bonificado da população para modelização não retirou informação relevante à carteira.

De igual forma constatou-se que o incumprimento no crédito concedido a colaboradores não é significativo e, como tal, esta segmentação não foi igualmente considerada na análise, para que não fosse igualmente factor de distorção dos dados e de enviesamento na análise.

Com base nas várias segmentações realizadas, a nossa população caracteriza-se por operações de Crédito Habitação realizadas a particulares, com empréstimos não bonificados e sem serem colaboradores da instituição.

Pode-se assim a partir deste ponto começar a isolar as operações que são inequivocamente de crédito a habitação, através da variável código do produto, das restantes operações.

Operações consideradas como Crédito Habitação:

- Crédito habitação Regime Geral;
- Crédito habitação com Carência de Capital.

-
- Obras de Beneficiação de Habitação;
 - Transferência de crédito habitação de outras Instituições

Outras operações não consideradas de Crédito Habitação:

- Pagamento de Sinal para Aquisição de Habitação;
- Crédito habitação para Aquisição de Propriedade Rurais;
- Aquisição de Terreno para Construção;
- Aquisição/Construção para Venda – Investimento.

Relativamente a estas últimas, houve alguma dificuldade em considerá-las como crédito à habitação, visto não terem como finalidade específica a compra de uma casa, mas sim finalidades diversas (terreno, investimento, etc), podendo não estar completamente associadas a verdadeiras operações de crédito à habitação, assim como, muitas delas possuem apenas garantias pessoais e não garantia hipotecária, o que lhes confere alguma proximidade de operação de crédito pessoal, mais do que ao crédito habitação hipotecário.

Seguidamente foram efectuadas outras filtragens aos dados disponíveis, nomeadamente naqueles que para além do número de observações não ser significativo, também não representavam a generalidade das operações, tendo sido por isso eliminadas, para não comprometerem os resultados.

Neste pressuposto, foram eliminadas:

- As operações com valor de crédito superior a 500.000 euros;
- As operações com valor inferior a 40.000 euros;
- As operações com prazo inferior a 13 meses;
- As operações sem garantia real;
- As operações cujos titulares possuíam idade superior a 70 anos no fim do contrato;
- As operações cujos titulares tinham idade inferior a 18 anos.

Com a realização desta segmentação e filtragem de dados procurámos então obter uma amostra homogénea, tão próxima quanto possível da realidade (da população inicial), e que nos permitisse identificar e estudar cada uma das variáveis relevantes para a determinação do incumprimento nas operações de crédito.

Resumidamente podemos considerar que o processo de desenvolvimento mencionado até aqui corresponde ao tratamento dos dados e construção da amostra para posterior análise estatística.

De forma esquemática apresentamos os passos realizados na recolha dos dados:

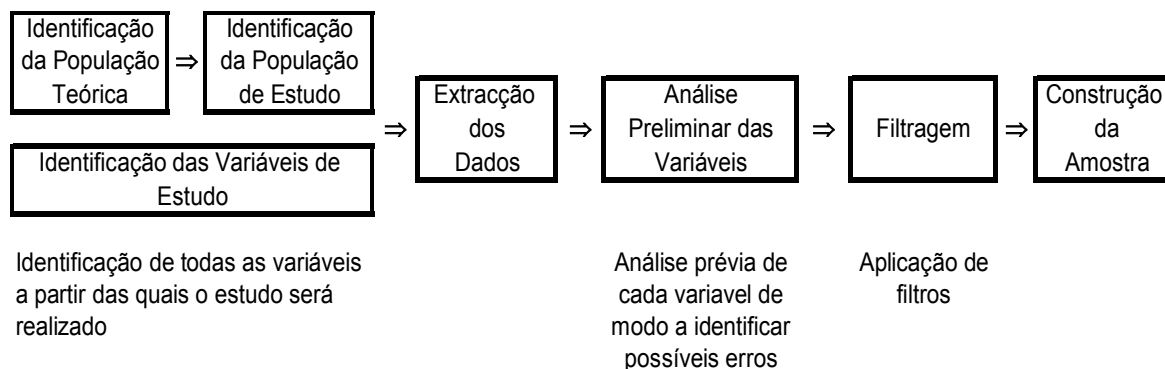


Figura 2 - Sistematização do processo de recolha de dados

De acordo com o esquema, podemos verificar que a partir da população total de dados, discriminou-se uma população de estudo, tendo sido a partir dela identificado um conjunto de variáveis, as quais iam sendo agrupadas, consoante as suas características do seguinte modo:

- Variáveis da operação – caracterizam a operação;
- Variáveis de comportamento da operação – determinam se as operações estão ou estiveram em incumprimento;
- Variáveis sócio demográficas – caracterizam os intervenientes num contexto social e demográfico;
- Variáveis de relação com a instituição – caracterizam a relação do cliente com o banco nos 12 meses anteriores à operação.

A partir da população inicial, e tendo por base, o conjunto de dados mais completos que se conseguiu obter, identificaram-se as variáveis que na nossa opinião (atendendo à informação recolhida na revisão bibliográfica), melhor caracterizariam o objectivo desta dissertação, pela sua importância na determinação do incumprimento.

Assim, consideramos como objecto de estudo nesta dissertação as seguintes variáveis:

- Valor da prestação;
- Prazo do empréstimo;
- Taxa de juro;
- Prazo do indexante
- Idade do cliente;
- Taxa de esforço;
- Rácio *loan-to-value*;
- Valor do Imóvel;
- Rendimento Anual.

De acordo com a revisão bibliográfica existem muitas outras variáveis, indicativas da ocorrência de incumprimento, algumas das quais, a sua inclusão neste estudo seriam de importância relevante, para a obtenção de resultados.

Neste tipo de variáveis podemos referir aquelas mais ligadas a caracteres sociais ou pessoais como desemprego, situação perante o emprego, estado civil, estado saúde, número de pessoas do agregado familiar.

Como referimos anteriormente, atendendo a que nem todas as operações possuíam devidamente carregados os dados, não nos foi possível obter dados fiáveis relativamente a este tipo de variáveis.

Ainda devido ao facto da população inicial não possuir todos os dados devidamente carregados (existem em suporte de papel, mas não em suporte informático), houve necessidade de realizar uma análise preliminar, que quer do ponto de vista da homogeneidade da informação, quer na exactidão ou validade da mesma, nos permitiu determinar a existência de “zeros” ou “missings” – observações sem informação - as quais iam sendo retiradas da base de dados.

O processo de filtragem consistiu em aplicar um conjunto de filtros a todas as variáveis de forma a eliminar da análise outras observações que apresentassem incongruências.

Concluído o processo de filtragem, a amostra de modelização obtida foi a seguinte:

- Totalidade das operações morosas da população de estudo;
- Doze operações não morosas por cada operação morosa.

3.2 Análise Descritiva

Na realização do processo estatístico utilizou-se numa primeira fase uma Análise Univariada, a qual constitui um processo estatístico que permite a análise e identificação das variáveis que, individualmente possuem um maior poder explicativo face à possibilidade de incumprimento.

As técnicas estatísticas da Análise Univariada têm como objectivo a descrição de cada uma das variáveis presentes no quadro de dados, recorrendo a medidas de síntese e a representações gráficas dos dados.

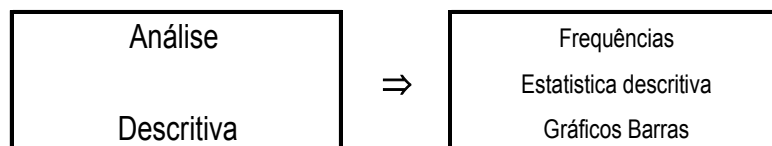


Figura 3 - Métodos de Análise Descritiva utilizados

Na fase da Análise Descritiva, são estudadas as características individuais de cada variável, de forma a se obter uma visão detalhada de cada uma, e a facilitar a escolha dos testes a aplicar sobre as mesmas.

Os estudos dependem assim do tipo de variável:

- Variáveis discretas – Frequências e/ou gráficos de Barras
- Variáveis discretas/contínuas – Estatísticas descritivas

3.3 Métodos utilizados na Análise de Default

A análise ao incumprimento de crédito é um tema que cada vez desperta mais interesse aos diversos agentes económicos, em particular às entidades financeiras. Foi a partir da década de 60 que começaram a surgir os primeiros estudos sobre a previsão de *default*. Na maioria dos trabalhos eram aplicadas técnicas estatísticas para obter modelos de previsão que permitissem dar indicações sobre a *probability of default* (Beaver, 1966).

Desde o trabalho pioneiro de Beaver (1966) diferentes métodos têm sido utilizados: análise discriminante, *logit*, *probit*, redes neuronais, indução de regras e árvores de decisão, algoritmos genéticos, conjuntos aproximados, entre outros modelos.

O objectivo deste trabalho não foi o de efectuar uma comparação entre os métodos mais utilizados, mas sim utilizar um método que fosse prático e com bons resultados.

Pela variada literatura recolhida, concluímos que a capacidade de previsão dos modelos é em geral similar e que a maioria dos investigadores utilizou a análise discriminante ou o *logit*.

A Análise Discriminante e a Regressão Logística, enquadram-se na classe de métodos estatísticos multivariados de dependência, uma vez que relacionam um conjunto de variáveis independentes com uma variável dependente categórica (Sharma, 1996; Hair *et al.*, 1998).

Para que se possam considerar válidos os resultados da Análise Discriminante é necessário a suposição de existir uma normalidade multivariada. Este método necessita que as variáveis independentes cumpram algumas condições, como serem normalmente distribuídas.

Ao contrário da análise discriminante, a regressão logística exige um número menor de pressupostos e que são menos rígidos (SPSS, 2003)

Para além destes aspectos, o resultado obtido na análise discriminante corresponde a uma pontuação, não traduzindo nenhum carácter probabilístico, o que leva à preferência pela utilização do Modelo Logístico (Ohlson, 1980).

O modelo *logit* possui também alguns inconvenientes. Mora (1996) refere para o modelo *logit*, algumas das limitações mencionadas na análise discriminante, nomeadamente a não consideração das probabilidades prévias e dos erros de classificação, sempre que se utilize como ponto de corte 0,5.

Segundo Ferrando & Blanco, (1998), uma das vantagens do modelo *logit*, prende-se com o facto de admitir que as variáveis independentes possam ser categóricas, permitindo que as variáveis explicativas não se reduzam unicamente a rácios económicos e financeiros ou variáveis métricas, possibilitando desta forma a utilização de informação não financeira ou qualitativa.

Hair (1998) refere também razões pelas quais a regressão Logística representa uma alternativa à análise discriminante, nomeadamente quando a variável dependente tem apenas duas categorias. Considera que a regressão logística é menos afectada pelas desigualdades variância/covariância e analisa facilmente variáveis categóricas independentes enquanto que na análise discriminante a utilização de variáveis *dummy* gera problemas com as igualdades variância/covariância.

Segundo Hosmer e Lemeshow (1989), a técnica de regressão logística tornou-se um método padrão de análise de regressão para variáveis medidas de forma dicotómica, sendo o seu objectivo determinar o melhor relacionamento entre a variável resposta (variável dependente) e um conjunto de variáveis explicativas, correspondendo o modelo final àquele que apresentar o melhor ajuste matemático e for mais fácil de se explicar.

Outra vantagem da regressão logística é a sua análise probabilística, uma vez que a regressão estima a probabilidade de ocorrer um certo evento, a partir de um conjunto de variáveis independentes ou explicativas.

Para Lo (1986) o modelo *logit* é mais robusto que a análise discriminante, já que é aplicável a outras distribuições que não a normal.

Dentro dos modelos de escolha discreta, a escolha de uma particular função $F(.)$ determinará o modelo específico com que se irá trabalhar. Assim, escolhendo-se a função de distribuição

normal estandardizada originará o modelo *probit*, escolhendo-se a distribuição logística originará o modelo *logit*.

Em comparação com o modelo *probit*, o *logit* tem representação e tratamento matemático mais simples, justificando a sua maior utilização prática.

À semelhança do que sucede com o *logit*, o modelo *probit* apresenta bons resultados quando a variável dependente é binária ou dicotómica (Borooah, 2002). Em termos práticos com o modelo *probit* obtêm-se as mesmas conclusões que com o *logit*, tendo a desvantagem dos coeficientes *probit* serem mais difíceis de interpretar, por isso a sua menor utilização. Uma vez que tem por base uma curva normal padrão, o *probit* não é recomendado quando existirem muitos casos assimétricos ou com outra distribuição que não a normal (Pampel, 2000).

3.1.1 Modelo de Regressão Logística - Logit

Atendendo a todos os aspectos mencionados no ponto anterior, considerámos que o modelo *logit*, seria o mais indicado para a realização do presente trabalho.

Este modelo baseia-se na função logística que surgiu em 1845 ligada a problemas de crescimento demográfico, (Meadows et al., 1972).

O *logit* obtêm-se a partir da regressão logística. Esta pode ser utilizada quando a variável dependente é binária ou dicotómica (Hosmer & Lemeshow, 1989). O procedimento que calcula os coeficientes logísticos compara a probabilidade de um evento suceder com a probabilidade de não suceder e designa-se por *odds ratio* (Hair et al., 1998).

Para Moore et al (1994) a análise *logit*, corresponde a um método que determina quais as variáveis que devem ser incluídas no modelo para que se possa prever adequadamente a

variável dependente categórica. Para o uso da análise logit as variáveis independentes e dependentes devem ser obrigatoriamente categóricas.

A regressão logística descreve um modelo de regressão linear, de modo a confirmar o valor da variável dependente entre 0 e 1, enquanto que as variáveis independentes podem variar continuamente.

A função logística está limitada em 0 e 1, de forma que previsões impossíveis não possam ocorrer. Ao prever o valor de uma variável nesta escala de 0 a 1, faz sentido ajustar uma curva em forma de S aos dados.

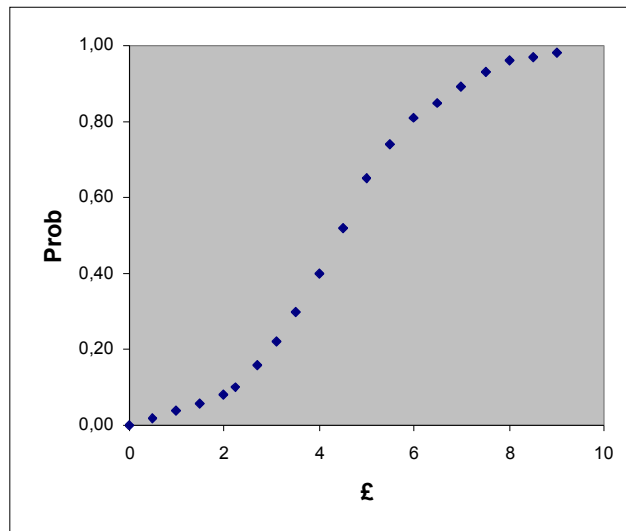


Figura 4 - Curva logística

Uma representação dos modelos, em que a variável dependente é dicotômica é a que pressupõe que a variável Y, (com Y = 1 ou Y = 0), é apenas uma indicação observável de uma variável não observável Y* (dita variável latente) tal que:

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

e em que é especificada uma regra de determinação de Y em função de Y*.

Essa regra é, tipicamente, da forma :

$$Y_i = \begin{cases} 1, & \text{Se } Y^* \geq 0 \\ 0, & \text{Se } Y^* < 0 \end{cases}$$

Onde a escolha da constante 0 como limiar da separação entre $Y = 0$ e $Y = 1$ é, essencialmente arbitrária.

Na classe de modelos caracterizada pelas relações das equações anteriores, podemos considerar que:

$$\text{Prob}(Y_i = 1) = \text{Prob}(Y_i^* > 0) =$$

$$= \text{Prob}(\beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i > 0) =$$

$$= \text{Prob}(u_i > -X_i \beta)$$

e por conseguinte,

$$\text{Prob}(Y_i = 0) = \text{Prob}(u_i \leq -X_i \beta).$$

Então, sendo u_i uma variável aleatória com função de distribuição $F(\cdot)$, vem que :

$$\text{Prob}(Y_i = 0) = F(-X_i \beta),$$

$$\text{Prob}(Y_i = 1) = 1 - F(-X_i \beta).$$

No caso do modelo *logit*, em que a forma funcional de $F(\cdot)$ é a que corresponde à distribuição logística, vem:

$$E(Y_i) = 0 [1 - F(X_i \beta)] + 1 F(X_i \beta),$$

ou seja,

$$E(Y_i) = F(X_i \beta).$$

Portanto,

$$E(Y) = \text{Prob}(Y = 1) = \frac{1}{1 + e^{-x}}$$

onde,

$$\text{Prob}(Y_i = 1) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}}$$

No fundo a regressão logística binária corresponde a uma regressão aplicada a uma variável dependente dicotómica, onde esta variável dependente não corresponde aos valores dos dados, mas representa a probabilidade do evento em análise vir a acontecer

Na regressão logística há um relacionamento linear com as variáveis independentes, mas esta linearidade acontece com a probabilidade do logaritmo e não na probabilidade original.

Supõe-se que o logaritmo da razão entre as probabilidades de ocorrência e não ocorrência do evento é linear. Como o objectivo do estudo é a probabilidade de acontecer determinado evento a equação logística pode ser transformada numa equação de probabilidade (Hair, 1999) assumindo assim a formula atrás apresentada.

Na regressão logística, uma vez que os erros do modelo seguem uma distribuição de Bernoulli, o método para estimar os coeficientes β_0, \dots, β_n é o método da máxima verossimilhança, o qual procura determinar os valores dos parâmetros β , que maximizam a probabilidade de se obter o conjunto dos dados observados (Hosmer e Lemeshow, 1998).

Na análise logística existem os métodos *enter*, *backward* e *forward*, os quais diferem entre si pela forma como seleccionam as variáveis que irão formar o modelo final.

A regressão logística minimiza o número de variáveis dependentes, de forma que o modelo final, possua mais aplicabilidade, uma vez que quantas mais variáveis existirem, mais dependente de dados fica o modelo.

O método *enter* caracteriza-se por incorporar ao modelo todas as variáveis em estudo (deverá assim ser utilizado quando se tem certeza de que todas as variáveis são necessárias para estimar os parâmetros do modelo).

Na regressão logística utiliza-se frequentemente a técnica *stepwise*, a qual corresponde a um processo de entrada ou retirada de variáveis ao modelo, baseado em critérios como por exemplo o teste Wald. Por exemplo, as variáveis com valores estatísticos (estatística Wald) mais elevados são os primeiros a ser incluídos no modelo.

O método *backward*, caracteriza-se por introduzir no modelo todas as variáveis e com o decorrer do processo vai sendo eliminada uma variável de cada vez, de forma que quando o processo de eliminação termina, temos definido o modelo final. Este processo ocorre assim por etapas, onde em cada etapa a variável de pior desempenho é eliminada, comparando-se o modelo completo com o modelo reduzido, resultante da eliminação de variáveis (Charnet *et al.*, 1999). O modelo resultante final corresponde assim àquele, com as variáveis que melhor caracterizam o objecto do estudo.

O método *forward*, caracteriza-se por considerar a variável de maior coeficiente de correlação na amostra, observado com a variável resposta. Em cada etapa uma nova variável pode ser adicionada de forma a construir o modelo final. Este processo ocorre assim por etapas, onde em cada etapa uma nova variável é acrescentada, comparando-se o modelo reduzido com o modelo em que uma nova variável é acrescentada. O processo decorre enquanto se verificar que com a inclusão de uma variável se obtém um modelo com melhor desempenho que o

anterior. Quando o processo de inclusão de variáveis termina, é porque não faz sentido a inclusão de mais nenhuma variável e as variáveis incluídas até essa data correspondem às que melhor caracterizam o objecto do estudo (Charnet *et al.*, 1999).

No presente trabalho vamos utilizar o método *Forward Stepwise* (Wald), através do qual, para obtenção do modelo final de regressão logística é necessário, no decorrer do processo, efectuar vários ajustes do modelo, ou seja, verificar se cada uma das variáveis obtidas nesse modelo está significativamente relacionada com a variável resposta do modelo. Esta análise é realizada através de testes de hipóteses estatísticos, que avaliam o modelo com e sem a variável.

Existem testes estatísticos para analisar a significância do modelo final, nomeadamente o teste Qui-quadrado e o teste Hosmer e Lemeshowd.

Estas duas medidas, em combinação, fornecem suporte para que se aceite o modelo de regressão logística como significativo. São estes testes que asseguram a evidência de significância estatística das variáveis

Primeiro realiza-se um teste Qui-quadrado para mudança no valor $-2 \text{ Log Likelihood}$ (-2 LL), do modelo base.

Neste caso, a significância de um modelo é dada pela mudança verificada em LL (Log Likelihood), estatística de verosimilhança, a qual é multiplicada por -2 $\{-2\text{LL}\}$, de forma a assumir um número positivo e a ter uma distribuição aproximada ao Qui-quadrado (χ^2), antes e depois de se introduzirem as variáveis no modelo. A referida diferença é apresentada sob a forma de Qui-quadrado, podendo observar-se a significância estatística desse e assim concluir-se acerca da significância do modelo.

Segundo Hair (1998), o ajuste geral do modelo pode ser avaliado utilizando-se a medida -2 LL . Se no modelo avaliado houver um decréscimo no valor -2LL comparado ao modelo base, existe melhoria no modelo, uma vez que menores valores da medida -2LL indicam o melhor ajuste do modelo.

Existem ainda três medidas de adequação do ajuste, as quais são comparáveis ao R^2 da regressão múltipla: A medida R^2 Cox e Snell, a medida Nagelkerke e medida Pseudo R^2 .

As medidas de adequação de ajuste comparam as probabilidades estimadas com as probabilidades observadas, para as quais valores mais altos indicam um ajuste melhor.

A medida R^2 Cox e Snell está limitada pelo facto de que não consegue alcançar o valor máximo de 1, de modo que Nagelkerke propôs uma modificação que tem o alcance de 0 para 1.

A medida Pseudo R^2 com base na melhoria do valor 2 LL.

O Pseudo R^2 é calculado como:

$$R^2_{\text{logn}} = \frac{2 \text{ LL base} - (-2 \text{ LL modelo})}{-2 \text{ LL base}}$$

Depois realiza-se a medida Hosmer e Lemeshow de ajuste geral, que tem um teste estatístico que indica se houve ou não diferença estatisticamente significativas entre as classificações observadas e previstas.

Mostra a não - significância indicando a ausência de diferença na distribuição de valores dependentes efectivos e previstos.

Esta é a medida final do ajuste do modelo e mede a correspondência dos valores efectivos e previstos da variável dependente.

Neste caso, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista.

Também neste teste um bom ajuste de modelo é indicado por um valor Qui-quadrado não significativo.

Finalmente temos as Matrizes de Classificação que mostram se as taxas de acerto são altas ou baixas para os casos correctamente classificados no modelo.

A validação do modelo de regressão logística pode ser obtida através da: criação de amostras de treino e validação (Hair, 1998).

Se as taxas de acertos da amostra de treino e da amostra de validação forem similares, pode-se concluir que o modelo tem suporte empírico no mesmo nível para explicar as variáveis dependentes.

Capítulo 3

4. Resultados

4.1 Análise Descritiva da Amostra

Após a recolha dos dados e constituição da amostra final, realizámos uma primeira análise, que consistiu numa análise univariada, na qual elaborámos um conjunto de quadros e gráficos, que espelham como é constituída a amostra e como se comporta cada uma das variáveis quando as operações se encontram em situação morosa (incumprimento) ou não morosa.

Relembramos que se considera uma operação morosa, aquela que se encontra com pagamentos em atraso há mais de 90 dias.

Com base nos dados do Quadro 1, podemos verificar que do total das 690 operações de crédito constantes da nossa amostra, 87 operações encontram-se em situação morosa, enquanto que 603 operações não estão em mora.

	Classificação do Crédito	
	Número Operações	Frequência
Operações morosas	87	0,126
Operações não morosas	603	0,874
Total Operações	690	1,000

Quadro 1 - Classificação do crédito da amostra - operações morosas e não morosas

A nossa amostra é constituída por 87,4 % de operações não morosas e 12,6% de operações em incumprimento. Temos assim, uma amostra com um número de observações de operações não morosas muito superior ao de operações morosas.

Embora não seja um dos objectivos deste trabalho, verificar se as operações não morosas, estão ou não em atraso, efectuou-se mesmo assim essa análise, donde podemos concluir (Quadro 2 e Gráfico 1), que das 603 operações não morosas, 487 estão em dia, mas existem 116 operações que se encontram com prestações em atraso, embora por um período inferior a 90 dias, e como tal, não tratadas neste trabalho, como estando em incumprimento.

	<i>Classificação do Crédito</i>	
	<i>Número Operações</i>	<i>Frequência</i>
Operações em dia	487	0,706
Operações com atraso Inf. 90 dias	116	0,168
Operações morosas (atraso sup 90 dias)	87	0,126
Total Operações	690	1,000

Quadro 2 - Classificação do crédito da amostra - Dias em atraso

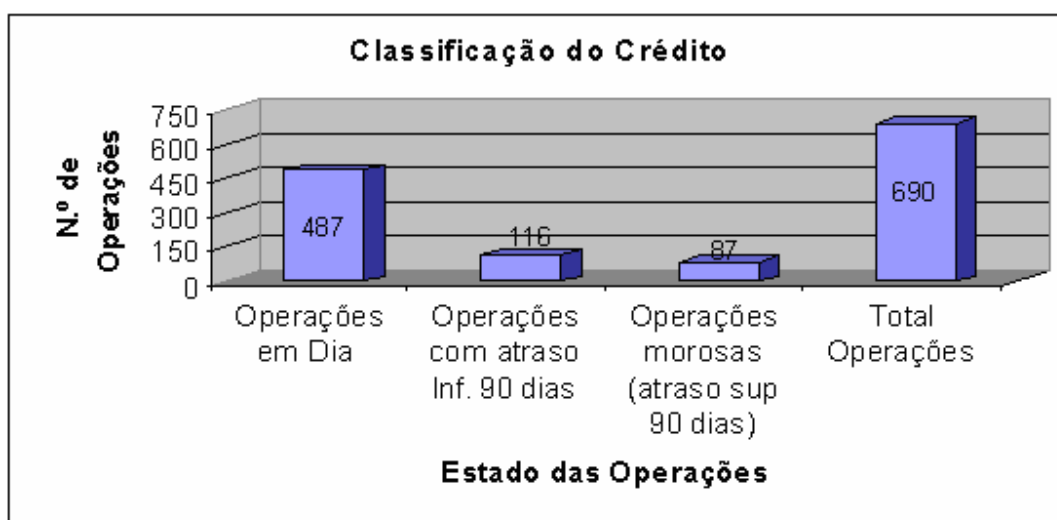


Gráfico 1 - Diversificação do crédito da amostra - Estado das operações

Para cada uma das variáveis em estudo efectuámos uma análise descritiva, de forma a se obter uma análise detalhada das suas características, bem como verificar os seus comportamentos quando as operações estão em situação de mora ou não.

Assim, nos pontos seguintes procuraremos apresentar, quer através de quadros, quer de gráficos, um resumo da análise efectuada para cada uma das variáveis.

4.1.1 Valor da Prestação

No Gráfico 2 podemos ver a representação por pontos dos dados da variável "Valor da Prestação", para as operações morosas e não morosas, as quais foram ordenadas de forma crescente. Devemos ter em atenção que a nossa amostra possui um maior número de dados para operações não morosas.

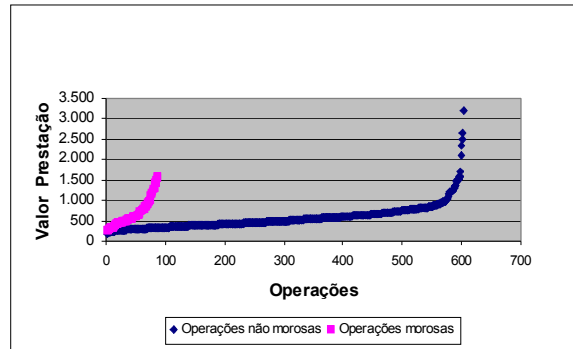


Gráfico 2 - Representação gráfica dos dados da amostra para variável "Valor da Prestação"

Com a realização da estatística descritiva, obtivemos os dados do Quadro 3, onde se verifica que a o valor médio das prestações, é maior no caso das operações que se encontram em situação morosa (676 euros), do que para as operações não morosas (567 euros).

Variável	Valor da Prestação	
	Operação não morosa	Operação morosa
Média	567 €	676 €
Mediana	497 €	562 €
Moda	814 €	#N/A
Desvio Padrão	300 €	331 €
Mínimo	192 €	268 €
Máximo	3.189 €	1.579 €

Quadro 3 - Estatística da variável "Valor da Prestação"

Através destes dados poderemos dizer, numa primeira análise que em média, as operações possuem uma prestação mais elevada nas operações que se encontram em incumprimento.

O desvio relativamente ao valor médio, ou seja, o desvio padrão, é quase igual nas duas situações, sendo neste caso o seu valor alto.

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude em qualquer uma das duas situações em análise.

No Gráfico 3, temos uma representação do comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

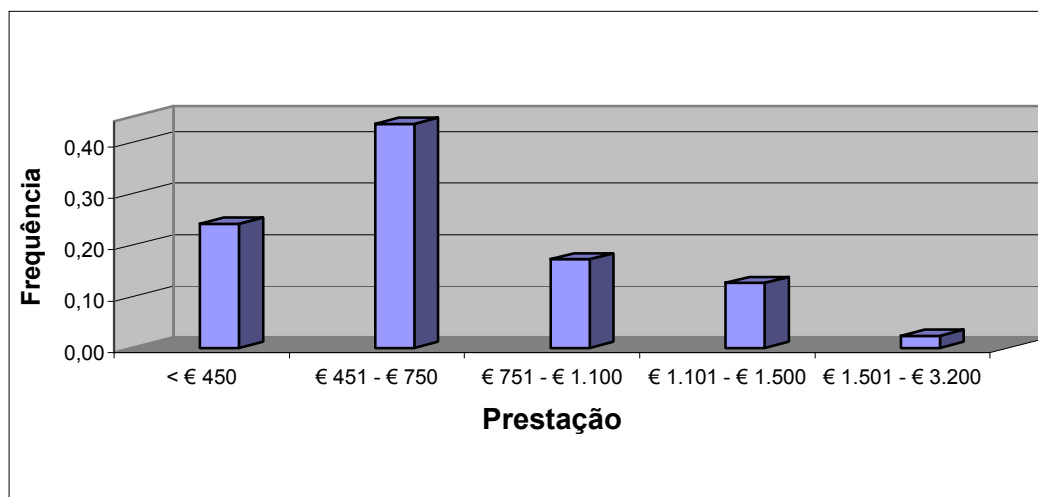


Gráfico 3 - Frequência de operações morosas para diferentes intervalos de "Valor de Prestação"

Pela análise do gráfico, constata-se que a frequência de incumprimento é maior quando o valor da prestação se situa entre 451 euros e 750 euros, seguindo-se as operações com menor valor de empréstimo (< 450 euros). A frequência de incumprimento é menor nas operações com prestações mais elevadas.

Estes resultados podem contudo estar associados ao facto do número de dados não ser igual em todos os intervalos.

Podemos verificar no Quadro 4, que em termos absolutos há mais operações de crédito com prestações entre 451€ e 750€, sendo neste intervalo que ocorre maior número de operações em mora.

Variável	Prestação		
	Operação morosa	Total Operações ^a	% Incumprimento
< € 450	21	264	8%
€ 451 - € 750	38	297	13%
€ 751 - € 1.100	15	91	16%
€ 1.101 - € 1.500	11	26	42%
€ 1.501 - € 3.200	2	12	17%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 4 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Valor da Prestação"

Em termos relativos a maior ocorrência de incumprimento ocorre quando o valor das prestações se situa no intervalo entre €1.101 e €1.500, com 42%, sendo menor nos empréstimos com as prestações mais baixas.

Estes resultados podem-nos apontar para uma realidade sócio-económica baixa, dado que as prestações mais baixas são as que reflectem o maior número de operações e consequentemente onde também há maior número de incumprimento, embora em termos percentuais o incumprimento ocorra numa gama de prestações mais elevadas.

4.1.2 Taxa de Juro

Seguindo os mesmos procedimentos que utilizámos no caso da variável anterior, podemos ver no Gráfico 4 uma representação da variável "Taxa de Juro", para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

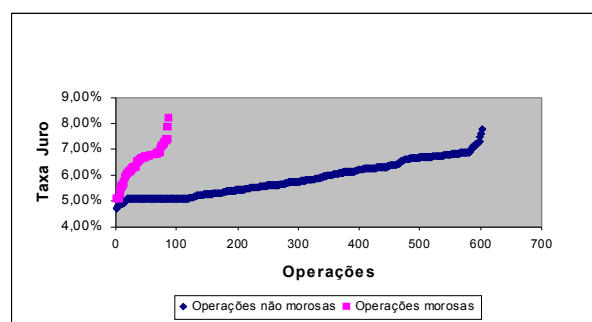


Gráfico 4 - Representação gráfica dos dados da amostra para a variável "Taxa de Juro"

O Quadro 5 apresenta um resumo da estatística descritiva, onde se pode verificar que o valor médio da taxa de juro é maior no caso das operações que se encontram em situação morosa (6,48%), comparativamente às não morosas (5,87%).

Variável	Taxa Juro	
	Operação não morosa	Operação morosa
Média	5,87%	6,48%
Mediana	5,75%	6,66%
Moda	5,10%	5,10%
Desvio Padrão	0,66%	0,64%
Mínimo	4,69%	5,10%
Máximo	7,80%	8,20%

Numa primeira análise, podemos referir, que em média, as operações possuem uma taxa de juro mais elevada nos empréstimos que se encontram numa situação de incumprimento.

O desvio relativamente ao valor médio (desvio padrão) é quase igual nas duas situações.

Quadro 5 - Estatística Descritiva da variável "Taxa de Juro"

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude, em qualquer uma das duas situações em análise.

O Gráfico 5, apresenta o comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

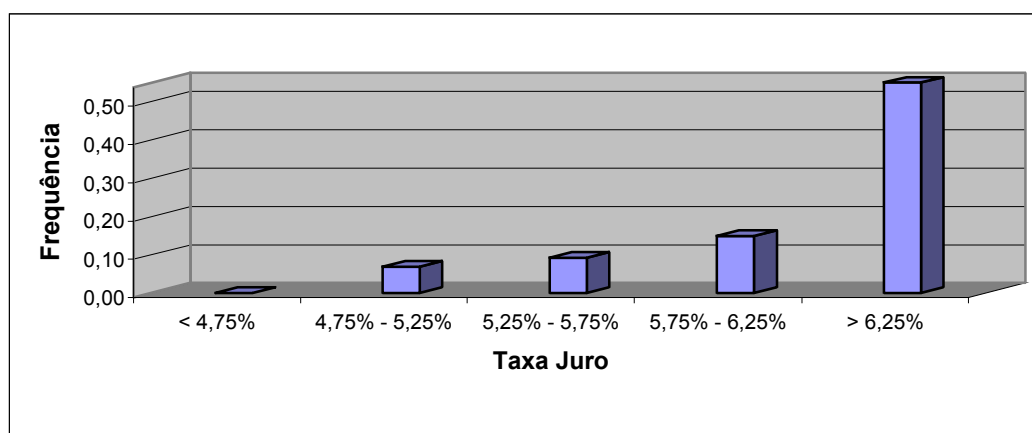


Gráfico 5 - Frequência de operações morosas para intervalos de "Taxa de Juro"

Pela análise do gráfico, podemos verificar que a frequência de incumprimento foi maior para operações com valores de taxa de juro mais elevadas, ou seja, é maior quando o valor da taxa se situa acima de 6,25%, seguindo-se as operações com taxas entre 5,75% e os 6,25%. Inversamente podemos referir que quanto menor o valor da taxa praticada nas operações, menor é a frequência de incumprimento registada.

Podemos verificar no Quadro 6, que em termos absolutos há mais operações de crédito com taxa de juro superior a 6,5%, sendo também neste intervalo que ocorre maior número de operações em mora.

Variável	Taxa Juro		
	Operação morosa	Total Operações ^a	% Incumprimento
Intervalo			
< 4,75%	0	1	0%
4,75% - 5,25%	6	161	4%
5,25% - 5,75%	8	154	5%
5,75% - 6,25%	13	131	10%
> 6,25%	60	243	25%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 6 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Taxa de Juro"

Em termos relativos verifica-se a mesma situação, em que a maior ocorrência de incumprimento ocorre quando a taxa de juro do empréstimo é superior a 6,25%, sendo menor nos empréstimos com as taxas mais baixas.

Estes dados vêm de certa forma confirmar alguns dos artigos referenciados na pesquisa bibliográfica, os quais apontam que o incumprimento está relacionado com as condições do empréstimo e tende a acontecer pelo aumento das taxas de juro.

4.1.3 Prazo do Empréstimo

Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no gráfico 6 uma representação da variável "Prazo do Empréstimo", para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

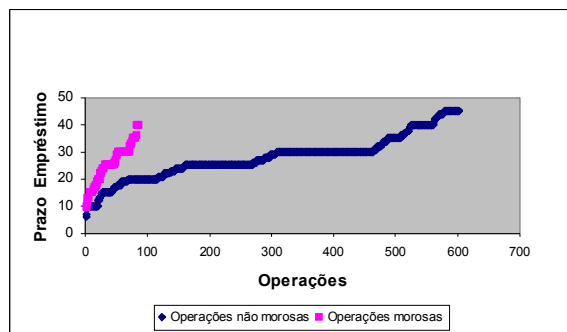


Gráfico 6 - - Representação gráfica dos dados da amostra para variável "Prazo de Empréstimo"

Após realização da estatística descritiva, obtivemos os dados do Quadro 7, onde se pode verificar que o prazo médio dos empréstimos é maior no caso das operações que se encontram em situação não morosa (28 anos), comparativamente às operações morosas (26 anos).

<i>Variável</i>	<i>Prazo Empréstimo (anos)</i>	
	<i>Operação não morosa</i>	<i>Operação morosa</i>
<i>Parâmetro</i>		
Média	28	26
Mediana	29	25
Moda	30	25
Desvio Padrão	8	7
Mínimo	6	10
Máximo	45	40

Através destes dados poderemos dizer, numa primeira análise que em média, as operações com menor duração, correspondem às operações que se encontram em incumprimento.

Quadro 7 - Estatística Descritiva da variável "Prazo de Empréstimo"

O desvio relativamente ao valor médio, ou seja o desvio padrão é quase igual nas duas situações.

Também aqui se verifica que entre os valores mínimos e máximos existe uma grande amplitude, em qualquer uma das duas situações em análise.

No Gráfico 7, apresentamos uma representação do comportamento da variável, dentro de certos intervalos, ou seja, a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

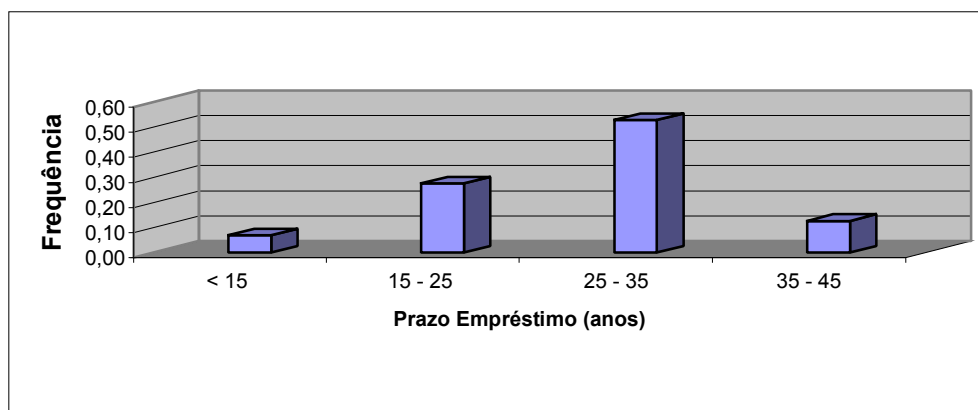


Gráfico 7 - Frequência de operações morosas para diferentes intervalos de "Prazo do Empréstimo"

Constata-se que a frequência de incumprimento é maior quando o prazo do empréstimo se situa entre os 25 e os 35 anos, seguindo-se as operações entre 15 e 25 anos. Verificamos que a frequência de incumprimento é mais baixa em empréstimos com menor duração (<15 anos) e em empréstimos de muito elevada duração (> 35 anos).

Estes resultados podem contudo estar associados ao facto do número de dados não ser igual em todos os intervalos.

Podemos confirmar pela análise do Quadro 8, que em termos absolutos há mais operações de crédito em empréstimos com um prazo entre 25 e 35 anos, sendo neste intervalo que ocorre maior número de operações em mora.

<i>Variável</i>	Prazo do Empréstimo (anos)		
	<i>Operação morosa</i>	<i>Total Operações^a</i>	<i>% Incumprimento</i>
<i>Intervalo</i>			
< 15	6	33	18%
15 - 25	24	157	15%
25 - 35	46	372	12%
35 - 45	11	128	9%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 8 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Prazo do Empréstimo"

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre quando o prazo dos empréstimos é inferior a 15 anos, com 18%, sendo menor nos empréstimos com prazos mais longos.

Estes dados confirmam o referido na bibliografia, na qual se refere que o aumento do prazo do empréstimo influencia na ocorrência de incumprimento, sendo menor quando se alargam os prazos dos empréstimos.

4.1.4 Idade do Cliente

Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no Gráfico 8 uma representação da variável “Idade do Cliente”, para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

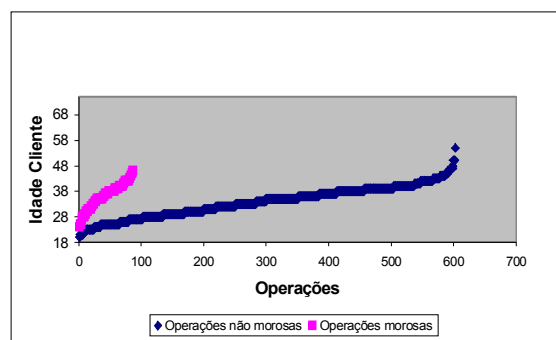


Gráfico 8 - Representação gráfica dos dados da amostra para variável "Idade do Cliente"

O Quadro 9 apresenta um resumo da estatística descritiva, onde se pode verificar que embora de forma pouco significativa, a idade média dos clientes, é maior no caso das operações que se encontram em situação morosa (36 anos), comparativamente às operações não morosas (34 anos).

Variável	Idade (anos)	
	Operação não morosa	Operação morosa
Média	34	36
Mediana	35	37
Moda	35	38
Desvio Padrão	6	5
Mínimo	20	24
Máximo	55	46

Através destes dados poderemos dizer, que em média, as operações que se encontram em situação de incumprimento são de clientes com maior idade do que os clientes com operações não morosas.

O desvio padrão é quase igual nas duas situações, sendo neste caso o seu valor alto.

Quadro 9 - Estatística da variável "Idade do Cliente"

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude em qualquer uma das duas situações em análise, contudo, as operações morosas encontram-se dentro de uma faixa de idades mais reduzida, ou seja varia entre os 24 e os 46 anos.

O Gráfico 9, apresenta o comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

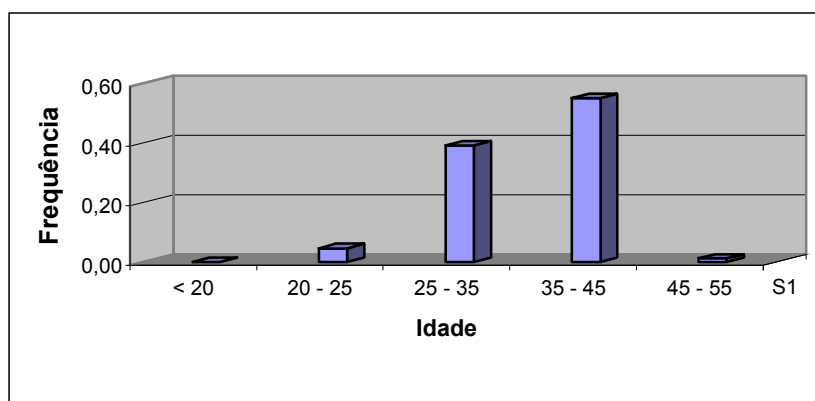


Gráfico 9 - Frequência de operações morosas para diferentes intervalos de "Idade do Cliente"

Pela análise do gráfico, podemos verificar que a frequência de incumprimento é maior quando em operação com clientes numa faixa etária entre os 35 e os 45 anos, seguindo-se as operações com clientes na faixa etária de 25 a 35 anos. A frequência de

incumprimento é menor nas operações com clientes de idades mais elevadas (> 45 anos) e em idades abaixo dos 25 anos.

Estes resultados podem contudo estar associados ao facto do número de dados não ser igual em todos os intervalos.

Pelos dados apresentados no Quadro 10, verifica-se que a maioria das operações foram contraídas por clientes na faixa dos 25 a 35 anos, logo seguido pela faixa dos 35 a 45 anos.

Em termos absolutos há mais operações realizadas com clientes na faixa etária dos 25 aos 35 anos, que traduz de certa forma, a necessidade dos casais mais jovens adquirirem habitação, sendo contudo na faixa dos 35 aos 45 anos que ocorre maior número de operações em mora.

<i>Variavel</i>	<i>Idade (anos)</i>		
	<i>Operação morosa</i>	<i>Total Operações^a</i>	<i>% Incumprimento</i>
<i>Intervalo</i>			
< 20	0	1	0%
20 - 25	4	68	6%
25 - 35	34	320	11%
35 - 45	48	287	17%
45 - 55	1	14	7%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 10 -- Quantificação das operações morosas (termos absolutos e relativos) para a variável "Idade do Cliente"

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre também na faixa etária dos 35 aos 45 anos, com 17%.

Estes dados vêm de certa forma confirmar alguns dos artigos referenciados na pesquisa bibliográfica, nomeadamente o estudo do OEC (2002), que afirma que o incumprimento tende a acontecer mais frequentemente na faixa etária dos 36 aos 45 anos.

4.1.5 Taxa de Esforço

A taxa de Esforço representa o quociente entre a prestação e a estimativa do rendimento disponível. Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no Gráfico 10 a sua representação, para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

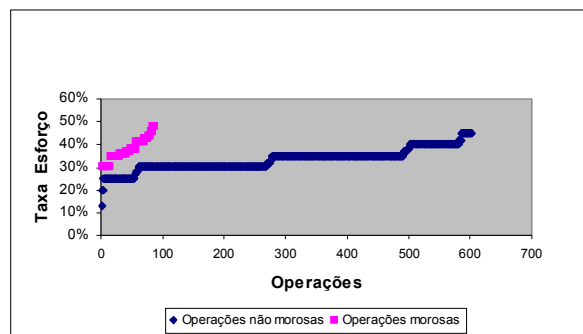


Gráfico 10 - Representação gráfica dos dados da amostra para a variável "Taxa de Esforço"

Com a realização de estatística descritiva obtivemos os dados do Quadro 11 onde se pode verificar que o valor médio da taxa de esforço obtido na amostra em análise, é maior no caso das operações que se encontram em situação morosa (38%), comparativamente à situação não morosa (33%).

Variável	Taxa Esforço	
	Operação não morosa	Operação morosa
Parâmetro		
Média	33%	38%
Mediana	35%	37%
Moda	35%	35%
Desvio Padrão	5%	5%
Minimo	13%	30%
Maximo	45%	48%

Quadro 11 - Estatística descritiva da variável "Taxa de Esforço"

Numa primeira análise podemos referir que em média, as operações que se encontram em situação morosa (incumprimento) possuem uma taxa de esforço mais elevada do que as operações não morosas.

O desvio padrão é quase igual nas duas situações.

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude, sendo muito maior no caso das operações não morosas.

O Gráfico 11, apresenta o comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

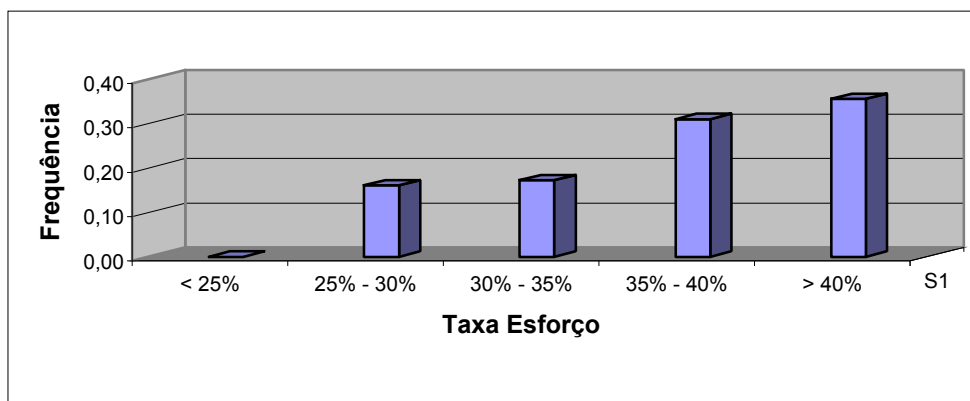


Gráfico 11 - Frequência de operações morosas para diferentes intervalos de "Taxa de Esforço"

Pela análise do gráfico, podemos verificar que a frequência de incumprimento foi maior quando o valor da taxa de esforço se situou acima dos 40%, seguindo-se as operações entre 35% e 40%. Verifica-se que a frequência de incumprimento foi menor nas operações com taxa de esforço mais baixas.

Pelos dados do Quadro 12, constata-se que em termos absolutos há mais operações de crédito realizadas a clientes com taxa de esforço entre 30% e 35%, contudo o maior número de operações em mora ocorra quando a taxa de esforço é superior a 40%.

<i>Variável</i>	<i>Taxa Esforço</i>		
	<i>Operação morosa</i>	<i>Total Operações^a</i>	<i>% Incumprimento</i>
<i>Intervalo</i>			
< 25%	0	55	0%
25% - 30%	14	228	6%
30% - 35%	15	238	6%
35% - 40%	27	117	23%
> 40%	31	52	60%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 12 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Taxa de Esforço"

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre nesta mesma situação, quando a taxa de esforço apresenta valores mais elevados, ou seja,

taxa de esforço superior a 40%, sendo menor nos empréstimos com taxas de esforço mais baixas.

Estes dados vêm de certa forma confirmar alguns dos artigos referenciados na pesquisa bibliográfica, os quais afirmam que o incumprimento, tende a acontecer quando as operações de crédito se realizam com taxas de esforço mais elevadas.

4.1.6 Loan - to - Value

Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no Gráfico 12 uma representação da variável “Loan to Value”, para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

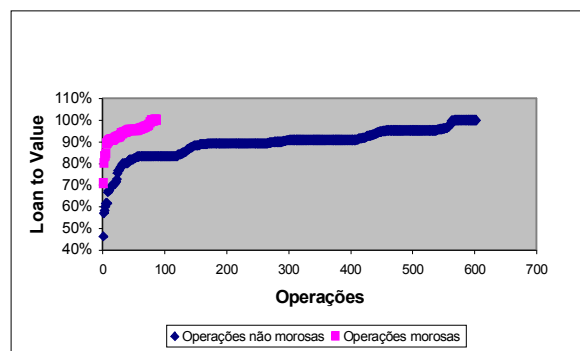


Gráfico 12 - Representação gráfica dos dados da amostra para a variável "Loan to Value"

O Quadro 13 apresenta um resumo da estatística descritiva onde se pode verificar que a o valor médio do *Loan to Value* (LVT), obtido na amostra em análise, é maior no caso das operações que se encontram em situação morosa (94%), comparativamente às operações em situação não morosa (90%).

Variável	Loan to value	
	Operação não morosa	Operação morosa
Parâmetro		
Média	90%	94%
Mediana	91%	95%
Moda	91%	95%
Desvio Padrão	7%	5%
Minimo	46%	71%
Maximo	100%	100%

Quadro 13 - Estatística descritiva da variável "Loan to Value"

Através destes dados poderemos referir que em média, as operações que se encontram em situação morosa (incumprimento) possuem um valor de *Loan to Value* mais elevada do que as operações não morosas.

O desvio relativamente ao valor médio, ou seja o desvio padrão é quase igual nas duas situações.

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude, sendo esta muito superior no caso das operações não morosas, a qual apresenta um valor mínimo muito mais baixo (46%) do que nas operações em incumprimento (71%).

No Gráfico 13, temos uma representação do comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

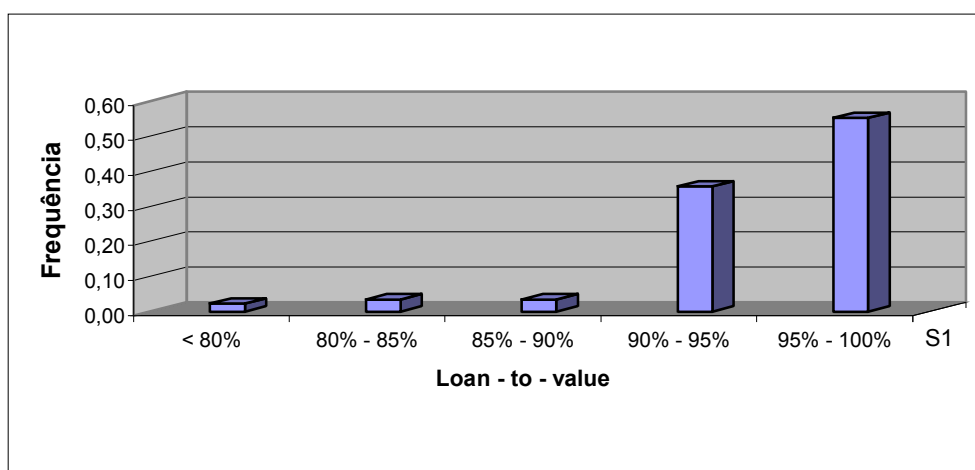


Gráfico 13 - Frequência de operações morosas para diferentes intervalos de “Loan to Value”

Pela análise do gráfico, podemos verificar que a frequência de incumprimento foi maior quando o valor de LTV se situa entre 95% e 100%, seguindo-se as operações com LTV entre 90% e 95%. A frequência de incumprimento foi menor nas operações com valores de LTV mais baixos.

O Quadro 14, mostra que em termos absolutos há mais operações de crédito com valores de Loan to Value entre 90% e 95%, contudo, o maior número de operações em mora ocorra nas operações de LTV entre 95% e 100%.

Variável	Loan - to - value		
	Operação morosa	Total Operações ^a	% Incumprimento
< 80%	2	42	5%
80% - 85%	3	94	3%
85% - 90%	3	160	2%
90% - 95%	31	198	16%
95% - 100%	48	196	24%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 14 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Loan to Value"

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre na mesma situação, quando o valor de LVT se situa no intervalo entre 95% e 100%.

Estes dados vêm de certa forma confirmar alguns dos artigos referenciados na pesquisa bibliográfica, os quais afirmam que o incumprimento tende a acontecer quando as operações de crédito se realizam com LVT mais elevados.

4.1.7 Valor do Imóvel

Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no Gráfico 14 uma representação da variável "", para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

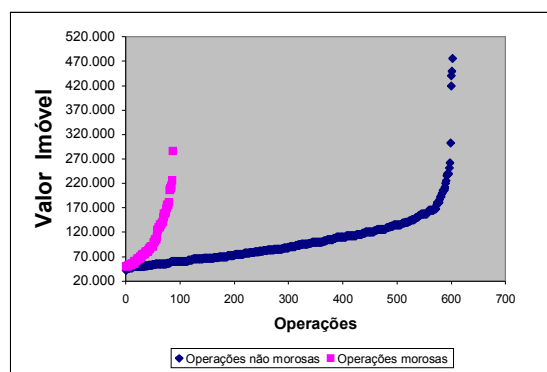


Gráfico 14 - Representação gráfica dos dados da amostra para a variável "Valor do Imóvel"

Com a realização da estatística descritiva, obtivemos os dados do Quadro 15 onde se pode verificar que o valor médio obtido na amostra em análise para o valor dos Imóveis, é maior no caso das operações que se encontram em situação morosa (104.117 euros), do que para as operações em situação não morosa (100.201 euros).

<i>Variável</i>	<i>Valor do Imóvel</i>	
	<i>Operação não morosa</i>	<i>Operação morosa</i>
<i>Parâmetro</i>		
Média	100.201 €	104.117 €
Mediana	89.250 €	85.000 €
Moda	60.000 €	55.000 €
Desvio Padrão	49.298 €	51.840 €
Minimo	41.899 €	50.000 €
Maximo	476.000 €	285.000 €

Quadro 15 -- Estatística Descritiva da variável "Valor do Imóvel"

Através destes dados, nomeadamente atendendo à média, poderemos dizer, numa primeira análise que em média, as operações que se encontram em situação morosa serviram para adquirir imóveis de valor mais elevado do que as operações não morosas, contudo a mediana apresenta-nos uma situação inversa.

O desvio relativamente ao valor médio, ou seja o desvio padrão é quase igual nas duas situações, sendo neste caso o seu valor alto.

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude, em qualquer uma das duas situações em análise.

No Gráfico 15, temos uma representação do comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

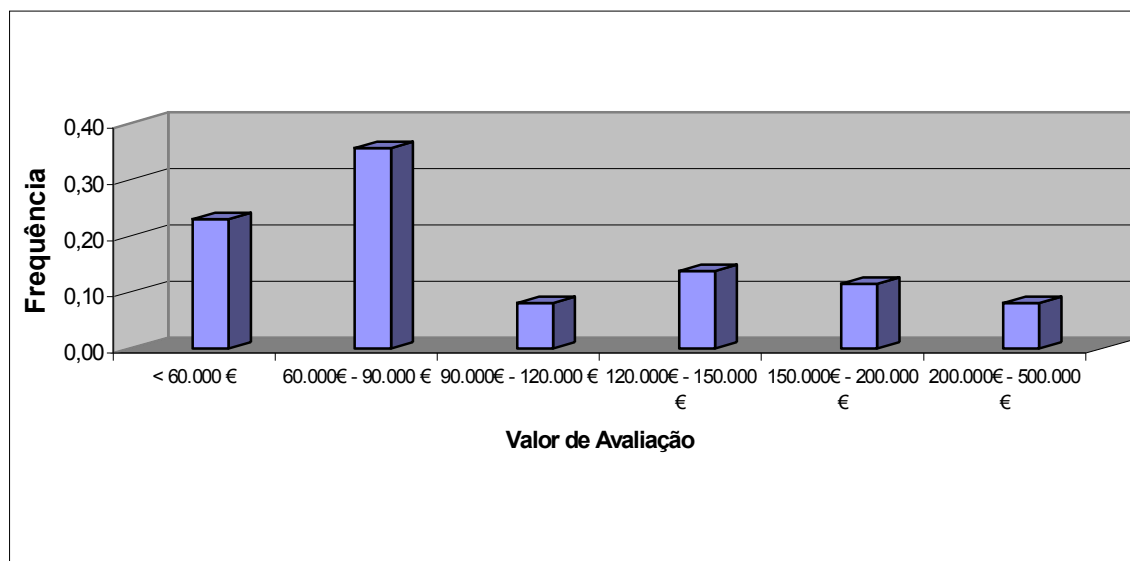


Gráfico 15 - Frequência de operações morosas para diferentes intervalos de “Valor de Avaliação”

Constata-se que a frequência de incumprimento foi maior nas operações que serviram para aquisição de imóveis com valor entre 60.000 euros e 90.000 euros, seguindo-se as operações para financiar imóveis com valor inferior a 60.000 euros.

Estes resultados podem contudo estar associados ao facto do número de dados não ser igual em todos os intervalos.

<i>Variável</i>	<i>Valor do Imóvel</i>		
	<i>Operação morosa</i>	<i>Total Operações^a</i>	<i>% Incumprimento</i>
<i>Intervalo</i>			
< 60.000 €	20	132	15%
60.000€ - 90.000 €	31	230	13%
90.000€ - 120.000 €	7	149	5%
120.000€ - 150.000 €	12	98	12%
150.000€ - 200.000 €	10	55	18%
200.000€ - 500.000 €	7	26	27%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 16 - Quantificação das operações morosas (termos absolutos e relativos) para a variável “Valor do Imóvel”

Verifica-se que a maioria das operações foram realizadas para aquisição de habitação com valor até €90.000, (onde se verificou a maior frequência de incumprimento), logo

seguido de operações para aquisição de habitações com valor entre €90.000 e €120.000, onde por sua vez, a frequência de incumprimento foi a mais baixa.

O Quadro 16, mostra que em termos absolutos há mais operações de crédito com imóveis entre 60.000 euros e 90.000 euros, sendo igualmente neste intervalo que ocorre maior número de operações em mora.

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre nas operações de crédito em que o valor dos imóveis adquiridos são mais elevados, ou seja no intervalo entre 200.000 e 500.000 euros, com 27%.

Estes resultados podem-nos apontar para uma realidade sócio-económica baixa, dado que a maioria das operações são realizadas para aquisição de imóveis de mais baixo valor.

3.1.8 Rendimento Anual do Agregado Familiar

Seguindo os mesmos procedimentos que utilizámos no caso das variáveis anteriores, podemos ver no Gráfico 16 uma representação da variável “Rendimento Anual do Agregado Familiar”, para as operações morosas e não morosas, as quais foram ordenadas de forma crescente.

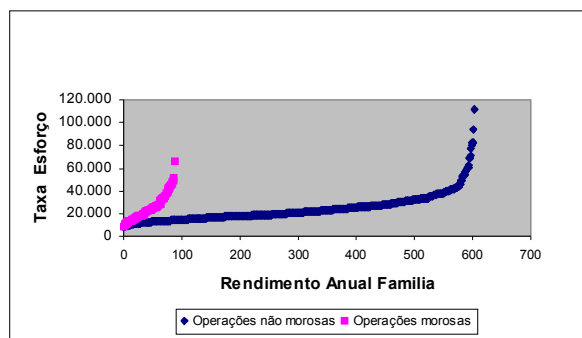


Gráfico 16 - Representação gráfica dos dados da amostra para variável “Rendimento Anual”

Após realização da estatística descritiva, obtivemos os dados do Quadro 17, onde se pode verificar que o valor médio, obtido na amostra em análise, para o rendimento anual, embora não seja significativo é maior no caso das operações que se encontram em situação morosa (€25.072), comparativamente às operações em situação não morosa (€24.030).

Variável	Rendimento Anual	
	Operação não morosa	Operação morosa
Parâmetro		
Média	24.030 €	25.072 €
Mediana	21.209 €	22.631 €
Moda	37.996 €	#N/A
Desvio Padrão	11.916 €	11.480 €
Mínimo	7.954 €	8.740 €
Máximo	111.625 €	66.185 €

Numa primeira análise podemos referir que em média, as operações que se encontram em situação morosa (incumprimento) são de clientes com rendimento mais elevada do que as operações não morosas.

Quadro 17 _ Estatística Descritiva da variável “Rendimento Anual do Agregado Familiar”

O desvio relativamente ao valor médio, ou seja o desvio padrão é quase igual nas duas situações, sendo neste caso o seu valor alto.

Verifica-se que entre os valores mínimos e máximos existe uma grande amplitude, sendo esta bastante maior no caso de operações não onerosas.

No Gráfico 17, temos uma representação do comportamento da variável, dentro de certos intervalos, ou seja a frequência de operações morosas em cada intervalo relativamente ao total das operações morosas.

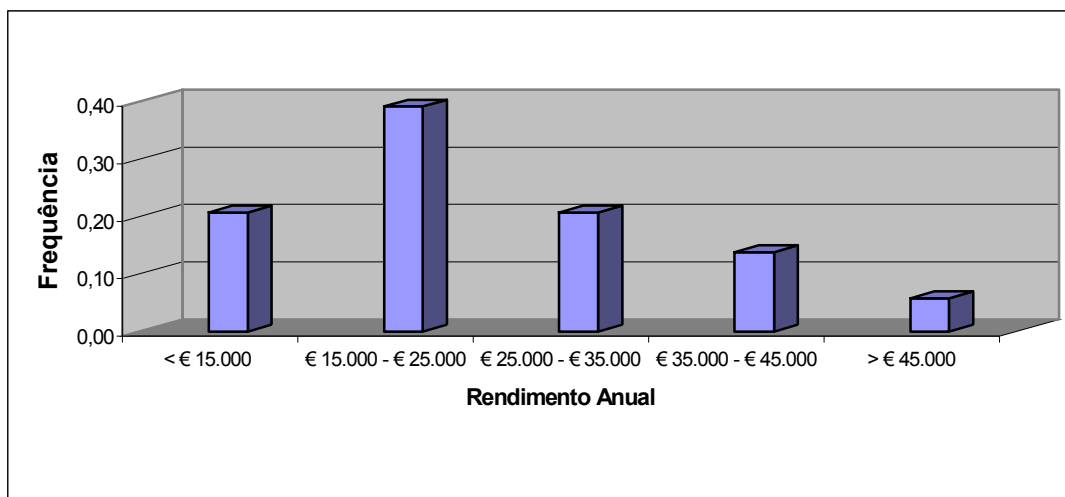


Gráfico 17 - Frequência de operações morosas para diferentes intervalos de “Rendimento Anual do Agregado Familiar”

Constata-se que a frequência de incumprimento foi maior quando o rendimento anual dos clientes se situa entre €15.000 e €25.000. Verifica-se que a partir de rendimentos superiores a €25.000, a frequência de incumprimento vai diminuindo à medida que o rendimento dos clientes vai aumentando.

Estes resultados podem contudo estar associados ao facto do número de dados não ser igual em todos os intervalos.

Podemos verificar pela análise do Quadro 18, que em termos absolutos há mais operações de crédito com clientes cujo rendimento anual se situa entre €15.000 e €25.000, sendo igualmente neste intervalo que ocorre maior número de operações em mora.

<i>Variável</i>	<i>Rendimento Anual</i>		
	<i>Operação morosa</i>	<i>Total Operações^a</i>	<i>% Incumprimento</i>
<i>Intervalo</i>			
< € 15.000	18	125	14%
€ 15.000 - € 25.000	34	325	10%
€ 25.000 - € 35.000	18	146	12%
€ 35.000 - € 45.000	12	63	19%
> € 45.000	5	31	16%
TOTAL	87	690	13%

a) Inclui operações morosas e não morosas

Quadro 18 - Quantificação das operações morosas (termos absolutos e relativos) para a variável "Rendimento Anual"

Em termos relativos verifica-se que a maior ocorrência de incumprimento ocorre quando o valor do rendimento se situa no intervalo entre €35.000 a €45.000.

Estes resultados podem-nos apontar para uma realidade sócio-económica baixa, dado que rendimentos mais baixos (até 25.000 euros) são os que reflectem o maior número de operações.

Estes dados, em valores absolutos, vêm de certa forma confirmar alguns dos artigos referenciados na pesquisa bibliográfica, os quais afirmam que o incumprimento tende a acontecer quando o rendimento das famílias se situa entre 750 a 1250 euros/mês, seguindo-se de 1.250 a 2.000 euros/mês, ou seja entre rendimentos anuais de 10.000 a 25.000 euros

4.1.9 Indexante

O Quadro 18, reflecte a forma como as operações estavam distribuídas conforme o indexante da taxa de juro, variava a 3 meses ou a 6 meses. Podemos constatar que do total de 690 operações, 181 estão indexadas a 3 meses e 509 indexadas a 6 meses.

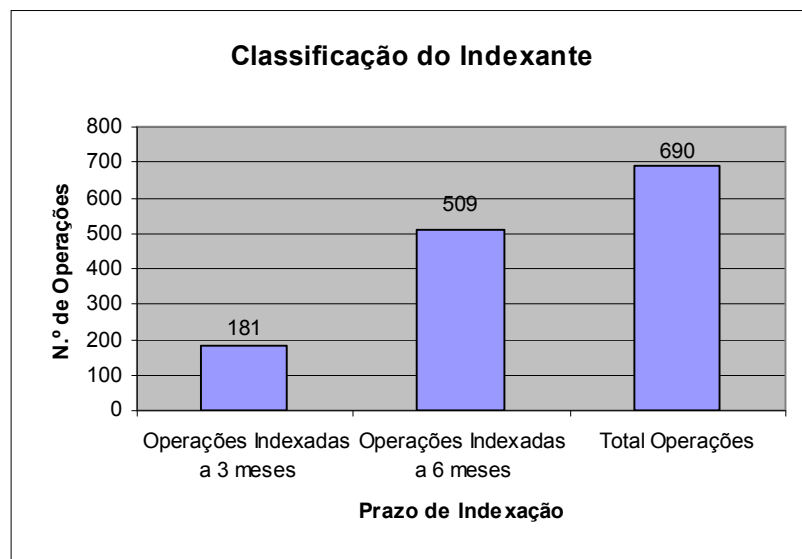


Gráfico 18 - Classificação do Indexante por prazo de indexação

Pela análise do próximo Quadro e respectivo gráfico, podemos verificar que a frequência de incumprimento foi muito maior nas operações indexadas a 6 meses (93,1%) do que nas operações indexadas a 3 meses (6,9%).

	<i>Indexante</i>		
	<i>Operações não morosas</i>	<i>Operações morosas</i>	<i>Frequência</i>
Operações Indexadas a 3 meses	175	6	0,069
Operações Indexadas a 6 meses	428	81	0,931
Total Operações	603	87	1,000

Quadro 19 - Classificação do Indexante - Frequência de incumprimento

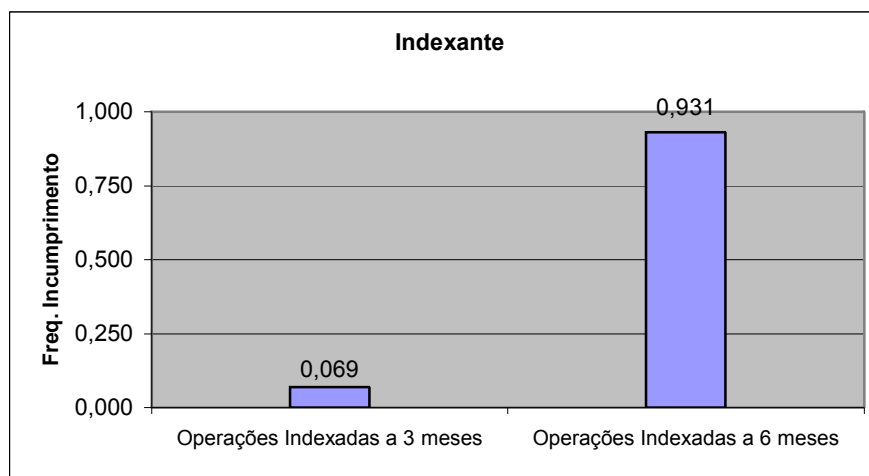


Gráfico 19 - Frequência de incumprimento consoante o indexante

4.2 Modelo de Regressão Logística - Logit

Após a análise Univariada das variáveis efectuamos uma análise Multivariada, onde utilizámos, como atrás referimos um modelo de escolha binária – Modelo de Regressão Logística (*Logit*).

Atendendo ao referido na bibliografia, na qual, embora com a utilização do modelo de análise discriminante, modelo probit ou logit, se obtenham resultados semelhantes, o modelo *logit* possui uma representação e tratamento matemático mais simples, bem como, mais facilidade de interpretação dos dados, optámos pela sua escolha na realização do presente trabalho.

As tabelas seguintes descrevem o *output* do programa SPSS na análise estatística do modelo de regressão *logit*, - no qual utilizámos o método stepwise forward (wald), para definir o modelo final. Este método tem por objectivo minimizar o número de variáveis e maximizar a precisão do modelo.

De acordo com o atrás mencionado, após segmentação e filtragem dos dados da população inicial obtivemos uma amostra composta pelas seguintes variáveis: Prestação, taxa de juro, prazo do empréstimo, indexante, idade do cliente, taxa de esforço, loan-to-value, valor do imóvel e rendimento anual, para as quais o presente trabalho tem por objectivo predizer as mais determinantes na ocorrência de incumprimento no crédito a habitação.

Um ponto importante na realização da regressão logística é a definição do ponto de corte. Quando a amostra reflecte a proporção existente na população, o ponto de corte que se utiliza corresponde à taxa verificada entre as duas situações em análise (neste caso operações morosas e não morosas).

Como no nosso caso, devido à necessidade de filtragem e eliminação de alguns dados para obtenção da amostra (devido como já se referiu anteriormente à base de dados não estar completa), esta não respeita a proporção existente na população. Nestas situações utiliza-se o ponto de corte de 0,5 que define probabilidades iguais para as duas situações em análise.

O modelo de regressão logística apresentado pelo programa SPSS inicia-se com a apresentação de um sumário da distribuição da amostra.

• **Distribuição da Amostra**

Pelo Quadro 20 podemos verificar que a amostra contém 690 dados. A não existência de *missings* deve-se ao facto de como vimos anteriormente, já terem sido eliminados aquando da obtenção da amostra.

		N.º	Percentagem
Dados seleccionados	Incluídos na análise	690	100%
	Missings	0	0
	Total	690	100%
	Casos não seleccionados	0	0%
	Total	690	100%

Quadro 20 - Case Processing Summary

• **Tabela de Classificação**

Seguidamente é apresentada a tabela de classificação (Quadro 21), onde considera o modelo com apenas a constante, ou seja, se arbitrariamente todos os empréstimos estivessem em situação não morosa, a taxa de acerto seria de 87,4%.

Observed			Predicted		
			Classificação do crédito		
			0	1	Percentage Correct
Step 0	Classificação do crédito	0	603	0	100
		1	87	0	0
		Overall Percentage			87,4

- a. Constant is included in the model.
 b. The cut value is 0,5

Quadro 21 - Classification Table a,b

• *Introdução de Variáveis*

Nos passos seguintes vão ser introduzidas sequencialmente variáveis no modelo de forma a se obter, o melhor modelo. A primeira variável a ser introduzida no modelo será aquela que tiver a estatística de pontuação mais alta, estatística de wald.

De acordo com o Quadro 22 podemos verificar que as variáveis foram seleccionadas em 4 etapas, ou seja a estatística Wald utilizou 4 passos até obter o modelo final. Pela observação das significâncias estatísticas do modelo, podemos constatar que o coeficiente é significativo em cada passo.

		Qui-quadrado	Graus liberdade	Significancia
Passo 1	Passo	60,576	1	0,000
	Bloco	60,576	1	0,000
	Modelo	60,576	1	0,000
Passo 2	Passo	73,710	1	0,000
	Bloco	134,286	2	0,000
	Modelo	134,286	2	0,000
Passo 3	Passo	40,528	1	0,000
	Bloco	174,813	3	0,000
	Modelo	174,813	3	0,000
Passo 4	Passo	13,437	1	0,000
	Bloco	188,250	4	0,000
	Modelo	188,250	4	0,000

Quadro 22 - Omnibus Tests of Model Coefficients

Pela análise da linha “Modelo”, podemos verificar que o modelo é estatisticamente significativo, rejeitando-se desta forma a hipótese nula, de que as variáveis independentes não são explicativas da variável dependente Y (operação morosa - incumprimento).

• *Avaliação do Ajuste Geral do Modelo*

O Quadro 23 apresenta três medidas para analisar o ajuste geral do modelo:

- - 2 Log Likelihood
- Cox & Snell (a qual está limitada visto que não obtém o valor 1)
- Nagelkerke (corrige a anterior e já apresenta um domínio de 0 a 1)

	- 2 Log Likelihood	Cox & Snell R ²	Nagelkerke R ²
Passo 1	462,278	0,084	0,158
Passo 2	388,568	0,177	0,333
Passo 3	348,041	0,224	0,421
Passo 4	334,604	0,239	0,449

Quadro 23 - Model Summary

Pela análise do quadro, podemos verificar que, à medida que foram introduzidas variáveis no modelo, houve uma redução no valor de -2 Log Likelihood, indicando uma melhoria no modelo.

O valor diminuiu de 462,27 no passo 1 para 334,60 no passo 3.

Esta diminuição determina que o modelo apresentado se encontra ajustado e o valor de 334,60 apresentado no último passo (passo 4) corresponde à medida geral em como o modelo se ajusta. Cada um dos valores de - 2 LL, apresentados em cada um dos passos, traduz o ajuste que se verifica de uma equação do modelo para outra com a introdução de mais uma variável.

De acordo com Hair (2005), um modelo bem ajustado possui um valor pequeno para - 2 LL, sendo o seu valor mínimo zero.

Contrariamente à medida - 2 Log Likelihood, no caso do R² de Cox & e do R² de Nagelkerke, a melhoria do modelo é traduzida pelo aumento do valor destas medidas à medida que se introduzem variáveis no modelo.

No Quadro podemos verificar que o valor R² de Cox & Snell varia de 0,084 para 0,239 e o valor R² de Nagelkerke (que corrige a anterior para poder apresentar um domínio de 0 a 1), varia de 0,158 para 0,449. Como se referiu anteriormente o aumento do valor destas medidas traduz uma melhoria no poder explicativo do modelo.

O R² de Nagelkerke no último passo aumentou 184% o poder de explicação do modelo obtido no passo 1.

Segundo Hair (2005), estas duas medidas comparam as probabilidades estimadas com as probabilidades observadas, sendo que valores mais altos significam um melhor ajuste do modelo.

Assim, podemos verificar pelo quadro que, à medida que foram introduzidas variáveis no modelo, houve um aumento do valor nestas duas medidas, ou seja à medida que foram introduzidas variáveis ao modelo, este foi-se ajustando melhor, de forma a apresentar valores mais altos.

Como referimos atrás a última medida de ajuste do modelo é o valor de Hosmer and Lemeshow Test, (Quadro 24) o qual mede a correspondência entre valores reais e os previstos da variável dependente. Possui um modelo estatístico que indica se ocorreram diferenças estatisticamente significativas entre as classificações observadas e previstas.

	Qui-quadrado	Graus de liberdade	Significancia
Passo 1	5,565	8	0,696
Passo 2	23,634	8	0,003
Passo 3	14,133	8	0,078
Passo 4	5,665	8	0,685

Quadro 24 - Hosmer and Lemeshow test

Nesta medida, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste do modelo é dado por um valor de Qui-quadrado não-significante.

Estas medidas combinadas (-2LL e Hosmer and Lemeshow Test), sugerem a aceitação do modelo do último passo como um modelo significativo de regressão logística.

Pelos resultados apresentados podemos concluir que o modelo obtido no passo quatro se encontra ajustado. Aceita-se este modelo de Regressão Logístico, como significativo.

• Avaliação da Precisão Estimativa do Modelo

Após se proceder à avaliação do Ajuste Geral do Modelo, segue-se uma avaliação da Precisão da Estimativa do Modelo, para a qual é utilizado o Método das Matrizes de Classificação.

O Quadro 25 constitui a matriz de classificação, a qual serve para avaliar a precisão de previsão do modelo.

Pela análise do quadro (matriz de classificação), podemos verificar a classificação dos casos preditos pelo modelo, comparando com os dados reais inerentes à variável dependente Y.

Observação	Preditos			
	Y			
	Operações não morosas 0	Operações morosas 1	Percentagem correcta	
Passo 1	0	599	4	99,34
	1	85	2	2,30
Percentagem geral				87,10
Passo 2	0	595	8	98,67
	1	62	25	28,74
Percentagem geral				89,86
Passo 3	0	588	15	97,51
	1	53	34	39,08
Percentagem geral				90,14
Passo 4	0	588	15	97,51
	1	53	34	39,08
Percentagem geral				90,14

a. The cut value is 0,5

Quadro 25 - Classification Table ^a

Verifica-se que o modelo apresenta uma melhor explicação no caso de operações não morosas, do que para as operações em mora. Explica correctamente 97,51% das operações não morosas e apenas 39,08% das operações morosas, ou seja explica correctamente 588 das 603 operações não morosas e apenas explica correctamente 34 das 87 operações em mora.

Esta menor explicação no caso das operações morosas pode estar associada ao facto de faltarem no modelo algumas variáveis, que podem ser consideradas de grande importância para a ocorrência de incumprimento. Como referimos a amostra obtida esteve condicionada à disponibilidade de dados existente na população, como tal não foi possível incluir algumas variáveis, que segundo a literatura são de grande importância na origem de incumprimento, nomeadamente factores sociais, como situação perante o emprego, vínculo laboral, estado de saúde.

Na aplicação do modelo deve-se ainda salientar o aspecto da amostra não possuir o mesmo número de observações para operações morosas e não morosas, principalmente o facto do número de operações morosas da amostra não representar a mesma proporção que a população, ou a mesma proporção do mercado.

Temos a noção que a existência de um maior número de situações morosas na nossa amostra relativamente às situações normais de mercado, ou da população, irão traduzir um β_1 real, inferior ao β_1 que se obterá no Modelo :

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

Em que : $\beta_1 \text{ real} < \beta_1 \text{ obtido na aplicação do nosso modelo.}$

A obtenção de um factor de correcção do β_1 do Modelo seria oportuno, mas careceria de um estudo muito aprofundado, o qual poderá ser objecto de uma tese futura.

No entanto considerando todo o modelo, verificamos que a taxa de acerto geral é de 90,14% ou seja, 90,14% das operações estão correctamente explicadas num dos dois grupos do modelo. Julgamos que no seu todo o modelo apresenta um bom poder explicativo, embora seja mais explicativo no caso das operações não morosas.

Se verificarmos o resultado obtido no modelo inicial, onde se considerava apenas a constante (Quadro 21), este tinha uma taxa de acerto de apenas 87,4%. Constata-se assim, que o modelo obtido no passo 4 aumenta a taxa de acerto de 87,4% para 90,14%.

Realça-se ainda que para casos em que as amostras em análise não possuem o mesmo tamanho (como é o nosso caso em que a amostra não possui o mesmo número de observações para operações morosas e não morosas), é conveniente identificar qual a percentagem mínima de classificação que é aceitável para o modelo.

Hair (2005) define um critério de proporcionalidade, o qual tem justamente em consideração a diferença de tamanho das amostras, com o qual vai identificar a percentagem mínima que é aceitável para o modelo.

Neste modelo possuímos 603 observações de operações não morosas ($Y=0$) e 87 operações morosas ($Y=1$).

A fórmula do critério proporcional é a seguinte:

$$CP = p^2 + (1 - p)^2$$

CP – Critério proporcional

p – proporção de operações não morosas

1 – p = proporção de operações morosas

Aplicando a fórmula obtemos:

$$CP = (0,873) + (0,126)^2$$

$$CP = 0,779$$

Como observámos a precisão de classificação do modelo 90,14% é superior ao critério proporcional obtido, daí considerarmos como aceitável o nosso modelo.

• **Validação do Modelo**

De acordo com o referido anteriormente, a validação do modelo de regressão logística pode ser obtida através da criação de amostras de treino e de validação (Hair *et al.*, 1998).

Como amostra de treino fomos utilizar uma amostra obtida aleatoriamente pelo programa SPSS, para a qual considerámos 300 operações (n.º escolhido a título de exemplo). A validação do modelo de regressão logística é obtida através da aplicação do modelo na amostra de validação.

Os quadros seguintes mostram a tabela de classificação com os resultados obtidos na amostra de treino (Quadro 27), os quais poderemos comparar com os resultados já apresentados no quadro 21, da amostra de validação (Quadro 26).

Observação	Preditos			Percentagem correcta
	Y			
	Operações não morosas 0	Operações morosas 1		
Passo 4	0	588	15	97,51
	1	53	34	39,08
Percentagem geral				90,14

a. The cut value is 0,5

Quadro 26 - Amostra de Validação (Classification Table a)

Observação	Preditos			Percentagem correcta
	Y			
	Operações não morosas 0	Operações morosas 1		
Passo 4	0	251	7	97,30
	1	22	20	47,60
Percentagem geral				90,30

a. The cut value is 0,5

Quadro 27 - Amostra de Treino (Classification Table a)

No geral podemos verificar as taxas de acerto na amostra de validação são quase idênticas às taxas de acerto na amostra de treino. Apenas no caso das operações morosas, estas são ligeiramente melhor explicadas na amostra de treino (explica correctamente 47,6%) do que na amostra de validação (explica correctamente 39,1% dos dados).

Verificando os dados gerais das duas amostras podemos concluir que o modelo de regressão logística possui um forte suporte empírico tanto na amostra de validação como na de treino para explicar as variáveis dependentes obtidas.

• Variáveis na Equação

O Quadro 28 representa as variáveis finais obtidas pela regressão logística, e os respectivos Betas, ou seja apresenta os vários parâmetros presentes na equação de regressão logística.

Podemos ver ainda o valor da estatística *Wald*, que identifica o quanto cada variável independente participa individualmente para a explicação da variável dependente. Observa-se pelo teste *Wald* que a significância de todos parâmetros foi aceitável.

Como podemos verificar o modelo final seleccionou quatro variáveis das nove inicialmente incluídas no modelo. Excluiu as variáveis: Indexante; Idade do Cliente; Valor do Imóvel; Rendimento Anual; e Prestação.

	B	S.E.	Wald	df	Sig.	Exp(B)	
Passo 1 ^a	Taxa Juro	1,371	0,192	50,785	1	0,000	3,940
	constante	-10,402	1,233	71,154	1	0,000	0,000
Passo 2 ^b	Taxa Juro	1,674	0,218	59,183	1	0,000	5,331
	Prazo Empréstimo	0,226	0,029	59,041	1	0,000	1,253
	constante	-19,489	2,008	102,128	1	0,000	0,000
Passo 3 ^c	Taxa Juro	1,725	0,228	57,469	1	0,000	5,612
	Prazo Empréstimo	0,220	0,031	49,884	1	0,000	1,246
	Taxa Esforço	0,172	0,031	30,443	1	0,000	1,187
	constante	-4,310	3,770	92,132	1	0,000	0,000
Passo 4 ^d	Taxa Juro	1,568	0,228	47,440	1	0,000	4,795
	Prazo Empréstimo	-0,073	0,020	12,894	1	0,000	0,930
	Taxa Esforço	0,256	0,034	57,517	1	0,000	1,291
	Loan to Value	0,186	0,032	33,103	1	0,000	1,205
	constante	-35,885	3,906	84,386	1	0,000	0,000

a. Variáveis introduzidas no passo 1:

b. Variáveis introduzidas no passo 2:

c. Variáveis introduzidas no passo 3:

d. Variáveis introduzidas no passo 4:

Quadro 28 - Variables in the Equation

Estes resultados demonstram que existe uma adequação do modelo para explicar o comportamento da variável dependente, onde as variáveis que resultaram no modelo final: Taxa de Juro; Prazo do Empréstimo; Taxa de Esforço; e Loan-to-Value; foram as que melhor explicaram o comportamento da variável dependente.

• **Aplicação da equação do Modelo e Cálculo da Probabilidade**

Assim, a partir do resultado dos parâmetros estimados pelo modelo *logit*, obtemos a seguinte equação final :

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

Onde, Y vem em função de Y^* :

$$Y_i = \begin{cases} 1, & \text{Se } Y^* \geq 0 \\ 0, & \text{Se } Y^* < 0 \end{cases}$$

$$Y^* = -35,885 + (1,568 \times \text{Taxa Juro}) + (-0,073 \times \text{Prazo Empréstimo}) \\ + (0,256 \times \text{Taxa Esforço}) + (0,186 \times \text{Loan-to-Value})$$

Aplicando os respectivos valores das variáveis no modelo de regressão linear, obtém-se para cada variável um valor de probabilidade de ocorrência de incumprimento, o qual como vimos anteriormente é definido por:

$$\text{Prob}(Y_j = 1) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki})}}$$

Aplicada à nossa amostra obteve-se uma probabilidade de incumprimento, em 49 operações do total das 690 operações da nossa amostra.

O Quadro 29 apresenta para vários intervalos, as probabilidades obtidas de ocorrência de incumprimento. Devemos ter em atenção que no desenvolvimento do nosso modelo efectuámos um ponto de corte de 0,5.

Intervalos Probabilidade	Morosas	Não Morosas	Total
0,00 a 0,25	0	577	577
0,25 a 0,50	0	64	64
0,50 a 0,75	29	0	29
0,75 a 1,00	20	0	20
Total	49	641	690

Quadro 29 - Intervalos de probabilidade de ocorrer incumprimento

Pela análise do quadro podemos verificar que existe entre 50 a 75% de probabilidade de 29 operações serem morosas e entre 75 e 100% de probabilidade de 20 operações serem morosas.

Discussão

Ao longo dos anos tem-se assistido a ciclos económicos mais ou menos longos, aos quais, depois de momentos de prosperidade ou euforia, surgem outros de mais estagnação ou declínio de algumas condições obtidas.

Nos últimos anos, verificou-se a nível mundial e particularmente em Portugal, uma conjuntura económica favorável ao investimento, através de incentivos aos jovens, taxas de juros significativamente mais baixas que no passado, com os bancos a aproveitar esta conjuntura para realizar negócio. Estes factores despoletaram um recurso ao crédito como até então não tinha acontecido.

Atingiu-se uma situação de endividamento das famílias elevado e preocupante, havendo a partir de certa altura, a noção que se as condições económicas do mercado fossem alteradas, poderiam advir problemas de incumprimento no pagamento das prestações por parte das famílias

Efectivamente, a conjuntura económica mudou e as condições de mercado actualmente são outras. A inflação atinge níveis mais elevados, houve aumento do desemprego, as taxas de juro são mais elevadas, e os próprios bancos ressentem-se da alteração das condições do mercado.

Com o atravessar deste ciclo económico o nível de incumprimento, como se esperava, aumentou consideravelmente.

De forma a prevenir estas situações e manter a estabilidade financeira dos bancos, surge o Novo Acordo de Basileia que define algumas regras prudenciais, relativamente à análise de risco de crédito por parte das entidades bancárias.

Tornou-se imperativo, os bancos passarem a definir metodologias de análise do risco de crédito. Estes podem utilizar métodos *standart* ou utilizar métodos de *ratings* internos, nos quais, são as próprias instituições que criam os seus próprios modelos internos, utilizando para tal as variáveis que entendam ser as mais explicativas de situações de risco.

O presente trabalho teve como objectivo, a partir de uma base de dados recolhida em alguns balcões de um banco nacional, aferir quais seriam as variáveis que mais poderiam contribuir para a ocorrência de incumprimento nas operações de crédito à habitação em curso.

A entidade bancária onde se realizou o trabalho, assumiu algumas limitações relativamente à sua base de dados, para as operações de crédito a habitação, estando em fase de conclusão uma plataforma informática onde cria a obrigatoriedade do carregamento destes campos para concretização das respectivas operações de crédito.

Assim, face aos dados que nos foram disponibilizados, deparámo-nos com alguns problemas, nomeadamente o facto da base de dados não estar completa, havendo muitos *missings* nas operações. Nem todos os dados relativos a cada operação estavam carregados informaticamente. Houve pois, necessidade de eliminar muitas operações, de forma a obter uma amostra que dispusesse da maior informação possível.

Algumas variáveis referidas na literatura como sendo de relevo, na justificação do incumprimento não puderam ser analisadas neste trabalho. Entre elas, salienta-se a situação profissional dos clientes (tipo de vínculo laboral, tempo de trabalho), estado de saúde, estado civil, tamanho do agregado familiar.

Perante a dificuldade em ter uma base de dados melhor, ou seja mais completa, obteve-se uma amostra, um pouco resumida e desproporcionada da população inicial e até das condições normais de mercado no que concerne ao número de operações morosas relativamente às não morosas. Na realidade a proporção de operações morosas / não morosas não reflecte a proporção existente na população.

Assim, a amostra traduz apenas a totalidade das observações para as quais foi possível obter o maior número de informação. Tentámos pois, reunir o máximo de informação, tendo por base variáveis que a literatura indicavam como sendo indicadoras de incumprimento.

A nossa amostra consistiu num conjunto de nove variáveis: Valor da Prestação; Prazo do Empréstimo; Taxa de Juro; Prazo do Indexante; Idade do Cliente; Taxa de Esforço; Rácio Loan-to-Value; Valor do Imóvel; Rendimento Anual do Agregado Familiar.

Para a realização deste estudo utilizou-se um Método de Regressão Logística – *Logit* - o qual, segundo a literatura, é de mais fácil aplicação e leitura dos resultados, traduzindo o *output* final numa probabilidade de ocorrência de incumprimento.

Algumas das limitações que encontramos na aplicação deste método prenderam-se essencialmente com a composição da amostra, por não representar proporcionalmente as operações morosas da população (daí ter-se utilizado no modelo um ponto de corte de 0,5 que assume probabilidades iguais para as duas situações, sendo o mais correcto a probabilidade estar de acordo com o realmente verificado na população), e pela falta de outros parâmetros explicativos da ocorrência de incumprimento.

Atendendo a estes aspectos, verificámos que o nosso modelo final, obtido no passo quatro, ou seja, com inclusão de quatro variáveis, explicava correctamente 97,51% das operações não morosas e apenas 39,08% das operações morosas.

Apesar destas limitações, o modelo no seu global mostrou-se ajustado e com elevado grau de precisão, com uma percentagem geral de 90,14%, ou seja, o modelo explica correctamente 90,14% do total dos dados da amostra.

Para garantir a fiabilidade destes dados, foram realizados testes de verificação do ajuste e da precisão do modelo, tendo-se sempre verificado que o modelo final obtido se encontrava ajustado.

Efectuou-se uma comparação entre a nossa amostra e uma amostra de treino (com apenas 300 dados), onde se concluiu que o modelo possuía suporte empírico, tanto na amostra de validação, como na de treino.

Das nove variáveis iniciais o modelo final aferiu quatro, como as mais explicativas da ocorrência de incumprimento na amostra: Taxa de Juro; Prazo do Empréstimo; Taxa de Esforço; Loan-to-value.

Obteve-se ainda a probabilidade de ocorrência de incumprimento, a qual foi apresentada sob a forma de intervalos.

Com a aplicação deste Modelo obtivemos assim (de acordo com a nossa amostra) o conjunto de variáveis explicativas da ocorrência de incumprimento e a probabilidade da sua ocorrência para cada operação.

Como exemplo de aplicabilidade desta metodologia poderíamos sugerir à luz destes dados (não descorando as limitações que existiram no desenvolvimento do processo), que uma nova

operação de crédito à habitação que surja na instituição, pode ser avaliada com base nestas variáveis e calculada a probabilidade de poder vir a ocorrer incumprimento.

Pensamos que esta metodologia pode, embora de forma simplista, ajudar as instituições de crédito a aferirem as variáveis que devem usar nos seus modelos de análise de risco de crédito.

Para que esta metodologia tenha uma melhor aplicabilidade é fundamental que em futuras investigações se obtenham bases de dados tão completas quanto possível e que retratem fidedignamente todas as operações da população, quer em termos do número de variáveis utilizadas, quer na proporção entre a população e a amostra.

Bibliografia

- Archer, W. et al (1999). "Determinants of multifamily mortgage default". Working paper. Federal Deposit Insurance Corporation. Junho.
- Archer, W., Ling, D. e McGill, G. (1996). "The Effect of Income and Collateral Constraints on Residential Mortgage Terminations". *Regional Science and Urban Economics*.
- Banco de Portugal (2005). *Relatório de Estabilidade Financeira 2004*. Lisboa.
- Banco de Portugal (2007). *Boletim Estatístico, Outubro de 2006*. Lisboa
- Bank of International Settlements (2005). "International Convergence of Capital Measurement and Capital Standards: A Revised Framework". Updated Version Novembro 2006.
- Beaver, W. (1966). "Financial Ratios as Predictors of Failure". *Journal of Accounting Research*.
- Bonfim, D. (2007). "Credit risk drivers: evaluating the contribution of firm level information and macroeconomic dynamics". Working Paper No. 7, Banco de Portugal.
- Bonfim, D. (2007). "Factores determinantes do risco de crédito: O contributo de características das empresas e da envolvente macroeconómica". Banco de Portugal. *Relatório de Estabilidade Financeira 2006*.
- Borooah, V. (2002). "Logit and Probit: Ordered and Multinomial Models". *Quantitative Applications in the Social Sciences, Series N.º 138*. Sage Publications: Thousand Oaks, California.
- Bradley, D., Nothaft, F. e Freund, J. (1998). *Financing Multifamily*.
- Calza e tal (2006). "Mortgage markets, collateral constraints, and monetary policy: do institutional factors matter?". *Centro de Estudos Financeiros da Universidade de Frankfurt*. Novembro.
- Câmara Municipal de Lisboa (2005). "Habitação e Mercado Imobiliário na Área Metropolitana de Lisboa". *Colecção de Estudos Urbanos – Lisboa XXI*. Junho.
- Cardoso, F. e Cunha, V. (2005). "Household wealth in Portugal: 1980-2004". Banco de Portugal, Working paper nº 4. Junho.
- Charnet, R. et al (1999). "Análise de modelos de regressão linear com aplicações". Campinas: Editora da Unicamp.
- Cunningham, D. e Hendershott, P. (1995). "Pricing FHA Mortgage Default Insurance". *Housing Finance Review*.
- Deng, Y. (1997). "Mortgage Termination: An Empirical Hazard Model with a Stochastic Term Structure". *The Journal of Real Estate Finance and Economics*.
- Dwight, M. e Quigley, J. (2007). "Housing Policy, Subprime Mortgage Policy, and the Federal Housing Administration". Agosto.

Elmer, P. e Seelig, S. (1999). "Insolvency, Trigger Events, and Consumer Risk Posture in the Theory of Single-Family Mortgage Default". *Journal of Housing Research* (forthcoming).

Elul, R. (2006). "Residential Mortgage Default". *Business Review*. Federal Bank of Philadelphia.

Farinha, L. (2003). "O efeito de algumas características demográficas e socioeconómicas sobre o endividamento das famílias". Banco de Portugal. *Boletim Económico*. Junho.

Farinha, L. (2004). "Grau de esforço associado ao endividamento das famílias: uma análise baseada em dados microeconómicos". Banco de Portugal. *Boletim Económico*. Setembro.

Farinha, L. e Noorali, S. (2005). "Endividamento e riqueza das famílias portuguesas". Banco de Portugal. *Relatório de Estabilidade Financeira 2004*.

Ferrando, M. e Blanco, F. (1998). "La Prevision del Fracaso Empresarial en la Comunidad valenciana: Aplicacion de los Modelos Discriminante y Logit" *Revista Española de Financiación y Contabilidad*.

Fitch Ratings (2005). "Portuguese Residential Default Model". Dezembro.

Fuinhas, J. (2003). "Canal do crédito, sobreendividamento e crise económica num contexto de União Monetária". *Revista de Gestão e Economia*. Universidade da beira Interior.

Genworth Financial Mortgage Insurance (2006). "Indicadores e Situação do Mercado Hipotecário Português". Junho.

Hair, J., Anderson, R., Tatham, R. e Black, W. (1998). "Multivariate Data Analysis". 5th ed. Prentice Hall.

Hair, J., Anderson, R., Tatham, R. e Black, W. (2005). "Análise multivariada de dados". 5ª. ed.. Porto Alegre: Bookman.

Hosmer, D. e Lemeshow S. (1989). "Applied Logistic Regression". John Wiley & Sons, Inc.: USA.

Hosmer, D., Lemeshow, S. (1998). "Applied Logistic Regression". New York: John Wiley & Sons, Inc.

Housing Finance International (2002). "Analysis of the Basel II Agreement and the Credit Risk in Residential Mortgage Loans". Novembro.

INE, Censos 1981.

INE, Censos 1991.

INE, Censos 2001.

Instituto Superior de Gestão Bancária (2005). "Análise Financeira de Bancos". Setembro.

Jiménez, G., e Saurina, J. (2006). "Credit cycles, credit risk and prudential regulation". *International Journal of Central Banking*. Junho.

Kau, J., Keenan, D. e Kim, T. (1994). "Default probabilities for mortgages. *Journal of Urban Economics*.

Lo, A. (1986). "Logit Versus Discriminant Analysis: a Specification Test and Application to Corporate Bankruptcy". *Journal of Econometrics*.

-
- Maroco, J. (2003). "Análise estatística com utilização do SPSS". Lisboa: Edições Sílabo.
- Marques, M. e Frade, C. (2003). "Uma sociedade aberta ao crédito". Subjudice, n.º 24.
- Marques, M. et al. (2000). "O endividamento dos consumidores". Coimbra: Almedina.
- Meadows et al (1972). "Limits to Growth". Potomac Associates: New York.
- Moore, W. et al (1994). "A Cross-Validity Comparison of Conjoint Analysis and Choice Models at Different Levels of Aggregation". Journal Marketing Letters. Vol. 9, nº 2. Abril.
- Mora, A. (1996). "El Contenido Informativo de los Datos Contables para las Decisiones de Inversion". Instituto de Contabilidad y Auditoría de Cuentas. Madrid.
- Observatório do Endividamento dos Consumidores (2002). "Crédito à Habitação – Endividamento e Incumprimento em Portugal". Centro de Estudos Sociais da Faculdade de Economia da Universidade de Coimbra. Fevereiro.
- Observatório do Endividamento dos Consumidores (2003). "Relatório de Actividades de 2002". Centro de Estudos Sociais da Faculdade de Economia da Universidade de Coimbra.
- Observatório do Endividamento dos Consumidores. "Mortgage Finance in Portugal: an overview".
- Ohlson, J. (1980). "Financial Ratios and the Probabilistic Prediction of Bankruptcy". Journal of Accounting Research.
- Pampel, F. (2000). "Logistic Regression: A Primer". Quantitative Applications in the Social Sciences, Series N.º 132. Thousand Oaks: Sage Publications.
- Pederzoli, C. e Torricelli, C. (2005). "Capital requirements and business cycle regimes: forward-looking modelling of default probabilities". Journal of Banking and Finance.
- Quigley, J. (1999). "A Decent Home: Housing Policy in Perspective". Berkeley Program on Housing and Urban Policy. Working Papers: Paper W99-007. Novembro.
- Sharma, S. (1996). "Applied multivariate techniques". New York: John Wiley & Sons.
- Yongheng, D., Quigley, J, e Order, R. (2006). "Mortgage Terminations, Heterogeneity, and the Exercise of Mortgage Options". Berkeley Program on Housing and Urban Policy. Working Paper Series 1007. Berkeley Program on Housing and Urban Policy.
- Zerbini, M. e Rocha, F. (2004). "Crédito ao consumidor: uma avaliação dos primeiros anos do Plano Real". Revista Nova Economia. Universidade Federal de Minas Gerais. Agosto.