

Altina Cecília Passo Katata

## **Fatores que Determinam a Falência das PME na Península Ibérica**



**Universidade do Minho**  
Escola de Economia e Gestão

Altina Cecília Passo Katata

## **Fatores que Determinam a Falência das PME na Península Ibérica**

Dissertação de Mestrado  
Mestrado em Finanças

Trabalho realizado sob a orientação da  
**Professora Doutora Sónia Maria da Silva Faria  
Nogueira da Silva**

## DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositórioUM da Universidade do Minho.

### *Licença concedida aos utilizadores deste trabalho*



Atribuição-NãoComercial-  
SemDerivações  
CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

## **Dedicatória**

Aos meus amados pais:  
Dario Daniel Katata (em Memória)  
E Aurora Nalindia Paulo Katata

## **Agradecimentos**

A DEUS, por estar sempre comigo durante todos os momentos da minha vida, nutrindo-me de fé, coragem, força e determinação para não desistir dos meus objetivos.

A minha mãe e a minha filha pelas orações, paciência, ajuda durante a estadia em braga e por terem aceitado embarcar comigo nessa desafiante aventura.

Ao meu amor Kelvin da Silva pelo apoio incondicional ao longo dessa caminhada, seria difícil sem ti.

Aos meus irmãos pelos concelhos, hospitalidade, torcida e principalmente por me mostrarem que manter a cabeça fria é sempre melhor.

A minha orientadora, Doutora Professora Sónia Maria da Silva Faria Nogueira da Silva por pacientemente ser a minha guia durante o desenvolvimento desta dissertação, que Deus te ilumine e te cuide sempre.

Aos meus queridos amigos Ana Isabel Sá e Manuel Bamba pelas ideias geniais.

Ao Professor Gilberto Loureiro e demais professores do Mestrado de Finanças da Universidade do Minho.

E por último aos restantes membros da família e amigos que direta ou indiretamente contribuíram para o sucesso dessa jornada.

Muito obrigada a todos.

## **Declaração de Integridade**

Declaro ter atuado com integridade na elaboração do presente trabalho acadêmico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducentes à sua elaboração.

Mais declaro que conheço e respeitei o Código de Conduta Ética da Universidade do Minho

## Resumo

Os modelos de previsão de falência mostram-se como ferramentas valiosas para a tomada de decisão, particularmente para as agências de notação de *rating*, gestores, banca, investidores, entre outros *stakeholders*.

Considerando que as micro, pequenas e médias Empresas (PME) são um dos pontos-chave para a prosperidade de qualquer economia, a presente dissertação tem como objetivo principal desenvolver um modelo de previsão que identifique os fatores determinantes de falência das PME da Península Ibérica. Para tal, foi usado um painel de dados da ORBIS Europe para um período compreendido entre 2012 e 2020, analisados com recurso à Análise Discriminante Múltipla e Análise Logística.

As variáveis que revelaram maior capacidade preditiva na deteção de falência são a Dimensão, a Solvabilidade, o Endividamento Total, o Rácio de Cobertura de Juros e o Índice de concentração industrial (HHI). Estas variáveis foram incluídas no modelo final de previsão de falência, que foi sujeito a diversos testes de robustez. O modelo final revelou um nível de precisão na deteção de falência que oscilou entre 81- 83%, até 3 anos antes do evento de falência.

**Palavras-chave:** Falência, Modelos de Previsão de Falência, Península Ibérica, PME.

## **Abstract**

Nowadays, financial distress prediction models are essential tools for decision making, particularly for rating agencies, managers, banks, and investors, among other stakeholders.

Considering that Micro, Small and Medium-sized Enterprises (SMEs) are one of the key points for the prosperity of any economy, the main objective of this dissertation is to develop a forecast model that identifies the determinants of SMEs' bankruptcy in the Iberian Peninsula. For this purpose, as used a panel of data from ORBIS Europe between 2012 and 2020, analyzed using Multiple Discriminant Analysis and Logistics Analysis.

The variables that showed the greatest predictive capacity in detecting financial distress are Size, Solvency, Total Leverage, Interest Coverage Ratio, and Industrial Concentration Index. These variables were included in the final prediction model, which was subjected to several robustness tests. The final model revealed an accuracy level in detecting bankruptcy that ranged between 81-83%, up to 3 years before the event.

**Keywords:** Bankruptcy, Bankruptcy Prediction Models, Iberian Peninsula, SMEs.



## Índice

<b>1. Introdução</b> .....	<b>1</b>
<b>2. Revisão da Literatura</b> .....	<b>4</b>
2.1 Definição de falência empresarial .....	4
2.2 Estudos/Modelos de previsão de Falência.....	6
2.3 Definição e Estudos de PME .....	13
2.3.1 Conceito de PME .....	13
2.3.2 Estudos de PME.....	14
2.4 Objetivos do Estudo.....	18
<b>3. Metodologia</b> .....	<b>20</b>
3.1 Análise Discriminante Múltipla .....	20
3.2 Modelos Logísticos.....	22
3.3 Propensity Score Matching .....	23
<b>4. Descrição da Amostra e das Variáveis</b> .....	<b>25</b>
4.1 Descrição da Amostra.....	25
4.2 Descrição das Variáveis .....	26
4.2.1 Variável dependente .....	26
4.2.2 Variáveis independentes .....	26
4.3 Estatísticas Descritivas .....	31
4.4 Análise da Correlação .....	38
4.5 Análise Univariada .....	40
<b>5. Análise Multivariada</b> .....	<b>42</b>
5.1 Análise Discriminante Multivariada.....	42
5.2 Análise Logística ( <i>Logit Analysis</i> ) .....	49
5.3 Propensity Score Matching .....	55
5.4 Apresentação dos Resultados.....	59
5.4.1 Modelo Logit com capacidade preditiva Um, Dois e Três Anos Antes do Evento de Falência.....	59
5.4.2 Modelo Final de Previsão de Falência .....	66
5.4.3 Teste de Robustez .....	69
<b>6. Conclusão</b> .....	<b>71</b>
<b>Referências Bibliográficas</b> .....	<b>73</b>
<b>Apêndice</b> .....	<b>76</b>

## Índice de Tabelas

Tabela 1 – Definição de PME .....	14
Tabela 2– Lista das variáveis independentes.....	27
Tabela 3 - Estatísticas Descritivas: Amostra Total .....	32
Tabela 4 - Estatísticas Descritivas por País (PME Espanholas vs PME Portuguesas) .....	33
Tabela 5- Estatística Descritiva por Status (PME Saudáveis vs PME Falidas) .....	34
Tabela 6 - Descrição da Amostra por Ano.....	35
Tabela 7 - Descrição da Amostra por País .....	35
Tabela 8 - Descrição da Amostra por CAE .....	36
Tabela 9 - Descrição da Amostra por NUTS II.....	37
Tabela 10 - Matriz de Correlação de Pearson .....	39
Tabela 11 - Teste às Médias e as Medianas por País e Status.....	41
Tabela 12 – Classificação das empresas conforme o seu status — Confusion Table.....	43
Tabela 13 - Erros tipos I e II decorrentes da classificação das empresas.....	43
Tabela 14 - Coeficientes da Função Discriminante.....	45
Tabela 15 - Coeficientes de Pearson — Structure Matrix .....	47
Tabela 16 - Análise ANOVA.....	48
Tabela 17 - Resultados antes do Processo de Standardização das Variáveis.....	52
Tabela 18 – Resultados após o Processo de Sstandardização das Variáveis.....	53
Tabela 19 – Estimação do Propensity Score.....	56
Tabela 20 – Resultados do teste realizado à qualidade do Matching .....	57
Tabela 21 — Análise Goodness-of-fit do Propensity Score.....	57
Tabela 22 — Resultados da Análise PSM-DID .....	57
Tabela 23 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência .....	60
Tabela 24 — Resultados da estimação do modelo Logit com os Erros-Padrão Robust e por Cluster.....	61
Tabela 25 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência, incluindo a variável Rentabilidade Operacional do Ativo e Rácio de Eficiência.....	63
Tabela 26 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência, controlando para Efeitos Temporais .....	65
Tabela 27 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência, controlando para efeitos temporais e usando PSM .....	66
Tabela 28 — Resultados da estimação do modelo Logit Final de Previsão de Falência, controlando para efeitos temporais.....	67
Tabela 29 — Resultados da estimação do teste Out-of-sample.....	69

## Índice de Gráficos

Gráfico 1 – Representação Gráfica dos Coeficientes <i>Standardized</i> .....	46
Gráfico 2 - ROC ( <i>Receiver Operating Characteristic Curve</i> ) .....	51
Gráfico 3 – Significância dos Coeficientes <i>Standardized</i> .....	54
Gráfico 4 – ROC ( <i>Receiver Operating Characteristic Curve</i> ): Modelo Logit Final .....	68
Gráfico 5 – Curva ROC do teste <i>Out-of-sample</i> .....	70

## **Lista de Abreviaturas**

ADM - Análise Discriminante Multivariada

ANOVA - *Analysis of Variance*

BCE - Banco Central Europeu

CAE - Classificação das Atividades Económicas

CIRE - Código da insolvência e da recuperação de empresas

EUA - Estados Unidos da América

FMI - Fundo Monetário Internacional

LR - *Likelihood Ratio*

PME - Pequenas e médias empresas

PS - *Propensity Score*

ROC - *Receiver Operating Characteristic Curve*

UE - União Europeia

PSM - *Propensity Score Matching*

PSM-DID - *Propensity Score Matching com Difference-in-Differences*

SMEs - *Small and Medium-sized Enterprises*

# 1. Introdução

As micro, pequenas e médias empresas (PME) são fundamentais para a economia de um país, abrangem uma diversidade de setores, fomentam o crescimento económico, criam emprego e riqueza. De acordo com o relatório *SME Country Fact Sheet* publicado em 2021 pela Comissão Europeia (European Commission, 2021), em 2020 as PME representavam 99,8% do tecido empresarial do setor não financeiro da União Europeia (UE), sendo as grandes empresas representadas pelos restantes 0,2%, onde as PME são responsáveis por 53,0% do valor acrescentado e 65,2% dos postos de trabalho deste setor (European Commission, 2021). Tal representatividade conduziu a UE a criar a lei das pequenas empresas (*Small Business Act- SAB*), cujo principal objetivo é garantir políticas e ações adequadas para as mesmas em prol de um ambiente inovador e propício para os negócios entre os Estados-Membros. Não obstante da existência da referida lei, é notório no referido relatório de 2021 da Comissão Europeia - *SME Country Fact Sheet* (European Commission, 2021) - que as PME na UE ainda têm muitos desafios pela frente, tais como: (i) aposta em competências, formação, inovação e digitalização; (ii) o acesso a financiamento continua a ser restrito, apesar de todos os processos adotados pela UE nos últimos anos para a sua melhoria; e (iii) definição de medidas de apoio para ajudar as PME a cumprir os regulamentos ambientais e energéticos em conformidade com o processo de transição verde da eu.

A crise financeira mundial de 2008, conhecida como crise do *Subprime* e cujo ponto alto culminou com a falência de um dos maiores bancos de investimento - o Lehman Brothers - começou nos Estados Unidos da América (EUA) e alastrou-se atingindo sobretudo o Sul da Europa, dando origem a um número considerável de eventos, entre os quais a degradação das condições económico-sociais das populações, o desemprego, o incumprimento por parte de Estados Soberanos e outros agentes económicos, conduzindo assim a um número acentuado de falências de empresas, particularmente de PME visto que elas apresentam um elevado nível de sinistralidade e, conseqüentemente, enfrentam maiores restrições no acesso ao financiamento em relação às grandes empresas, principalmente em tempos de crise, revelando menor capacidade para aumentar o seu investimento no período que se

seguiu à crise financeira (European Central Bank, 2013). Ora, para evitar eventos de crédito por parte dos Estados Soberanos da UE, foram implementados programas de austeridade como condição de resgate para alguns países como a Grécia, Portugal, Irlanda e Espanha que acabaram por ser intervencionados pela União Europeia (UE), o Fundo Monetário Internacional (FMI) e o Banco Central Europeu (BCE), que ficou conhecida como 'Troika'. Na atual conjuntura em que o mundo vive devido à pandemia global causada pela Covid-19 e pelas medidas adotadas para a contenção da transmissão desse vírus, assim como o conflito armado Rússia-Ucrânia, esses eventos são inevitáveis e acredita-se que podem vir a gerar uma recessão ainda pior do que a vivida anteriormente.

Ciclos económicos recessivos exigem a melhoria do acesso a financiamento para a atividade empresarial, evitando dessa forma a deterioração da situação financeira das empresas e fomentando a recuperação rápida desse ciclo. Seguindo essa abordagem, a UE aprovou um pacote de apoio financeiro para a sua recuperação que tem as PME como um dos principais beneficiários, visto que por elas serem empresas com um nível de complexidade baixo, podem se facilmente adaptadas às mudanças das condições económicas (Altman & Sabato, 2007). Apesar de essa ajuda evitar uma desaceleração muito mais severa, a sua eficácia varia entre Estados-Membros. Nesse contexto de endividamento, se o apoio for retirado e os fundos disponibilizados não alcançarem o efeito desejado como consequência de uma possível gestão ineficiente, a falência pode ser inevitável para muitos agentes económicos.

Dada a importância das PME para o desenvolvimento económico de um país, bem como as dificuldades que elas encontram no acesso a financiamento e sendo um dos pontos-chave para a recuperação da economia europeia, o presente trabalho tem como objetivo identificar e analisar variáveis capazes de explicar a probabilidade de falência das PME na Península Ibérica<sup>1</sup> durante o período de 2012-2020, através do recurso a informação contabilística e financeira, usando diferentes metodologias como a análise univariada, a análise discriminante multivariada (ADM) e a análise logística para desenvolver um modelo empírico para PME visto que elas diferem das grandes empresas do ponto de vista do risco de falência em conformidade com estudos existentes (p.e. Altman e Sabato, 2007, Gupta et. al, 2014).

Com ou sem cenário de crise macroeconómica, a probabilidade de

---

<sup>1</sup> Espanha, Portugal.

incumprimento pode ser uma realidade para um elevado número de PME. Como consequência, os agentes económicos – Estado, clientes, fornecedores, etc. - estão cada vez mais atentos a questões relacionadas com o risco de incumprimento para reagir atempadamente, evitando dessa forma potenciais perdas. Cada vez que um evento de falência empresarial ocorre, é importante detetar as suas causas como forma de prever eventos futuros da mesma natureza.

Os modelos de previsão de falência são um tópico para o qual existe uma extensa literatura e muitos dos modelos desenvolvidos inicialmente focaram-se em informação contabilística - Balanço, Demonstração de resultados, etc. – como Beaver (1966), e/ou informação de mercado (p.e., Altman, 1968; Ohlson, 1980). Com o Acordo de Basileia II, surgiram alguns estudos direccionados a PME, cujo objetivo era abordar o problema dos possíveis efeitos do Acordo de Basileia II sobre os requisitos de capital dos bancos, mas o problema de prever uma situação de falência especificamente para as PME foi apenas brevemente considerado (Altman & Sabato, 2007). Ora, desde então é cada vez mais visível o interesse de investigadores, académicos, economistas (inclusive de instituições financeiras), em estudar os determinantes de falência para esse segmento de empresas. Neste trabalho, propomos preencher esse *gap* no que respeita ao contexto da Península Ibérica.

O presente trabalho encontra-se estruturada da seguinte forma: No Capítulo 2 é apresentada a revisão da literatura com destaque para os conceitos diferenciados de falência e alguns dos métodos desenvolvidos para previsão de falências para grandes empresas e PME; no Capítulo 3 será abordada a metodologia a aplicar na análise empírica; o Capítulo 4 apresenta-se a descrição das variáveis, da amostra e a sua origem; no Capítulo 5 são apresentadas os resultados empíricos, com destaque para o modelo de previsão de falência e a referida análise dos resultados; no Capítulo 6 são apresentadas as conclusões deste estudo.

## 2. Revisão da Literatura

### 2.1 Definição de falência empresarial

Como ponto de partida é importante percebermos o conceito de falência porque apesar de genericamente associar-se o termo a dificuldades que uma empresa apresenta para fazer face as suas obrigações financeiras, com fornecedores, credores, funcionários e outros, não deixa de ser um facto de que na literatura são várias as definições existentes em torno do conceito, sendo usual a utilização dos seguintes termos: *failure*, *insolvency*, *default* e *bankruptcy*, que apesar de serem usadas sem qualquer distinção, formalmente são diferentes (Altman & Hotchkiss, 1993). Os autores definiram *failure* como o estado em que a taxa de retorno dos investidores na empresa é inferior às taxas de investimentos idênticos, quando a empresa apresenta insuficiência para pagar as dívidas e o seu retorno médio está sempre abaixo do seu custo de capital, porém esse estado não implica a extinção da empresa. Quanto a *insolvency*, consideraram que ocorre quando a empresa não tem liquidez para cumprir as suas obrigações de curto prazo. Já em relação ao *default*, os autores mencionam que este acontece quando a empresa viola um acordo contratual com um credor, que pode dar lugar a uma ação legal contra a empresa. E finalmente definiram *bankruptcy* como um procedimento legal para a extinção da empresa.

Ao longo dos anos a definição de falência variou entre estudos, mas não obstante das divergências, frequentemente existem pontos comuns nas variadas definições. De acordo com Beaver (1966), falência é definida como a incapacidade de uma empresa pagar as suas obrigações financeiras na data do seu vencimento. Associando ainda a falência operacional de uma empresa a eventos como: incumprimento no pagamento de títulos, contas bancárias a descoberto ou o não pagamento de dividendos. Enquanto Altman (1968) adotou a definição segundo o capítulo X da lei de falências dos EUA, que se refere as empresas falidas como as que estão legalmente falidas, em situação de liquidação ou ainda em processo de reestruturação. Por sua vez, Ohlson (1980) optou também por uma definição jurídica, ou seja, empresas declaradas falidas ou que já deram início ao pedido de falência, com base no capítulo X e capítulo XI da lei de falências dos EUA. Sendo que, os processos do capítulo XI são mais frequentes e de curta duração e mais bem-sucedidos do que os processos do capítulo



X pela sua complexidade. Blum (1974) definiu falência referindo-se a indícios que evidenciem a incapacidade de a empresa em pagar dívidas e, conseqüentemente, a reestruturação da referida dívida, ou ainda a petição de um processo de falência. (Altman & Hotchkiss, 1993).

Por sua vez, Neves (1997) no seu estudo não definiu falência como um evento, mas sim como o declínio da situação económica da empresa, que segue um determinado processo dividido em sete fases. Na primeira fase observam-se regularmente indícios que afetam negativamente o desempenho económico da empresa, e subsequentemente numa segunda fase, quando a tesouraria líquida atinge valores negativos, a empresa tende a apresentar uma situação de pouca ou mesmo ausência de liquidez. Na terceira fase, a empresa começa a apresentar incumprimento das suas obrigações, fase essa que reflete o definido no artigo 3º do CIRE (Código da Insolvência e da Recuperação de Empresas) em Portugal.

Na quarta fase, o fundo de maneo torna-se negativo e os incumprimentos acumulam-se, dando lugar a insolvência parcial na quinta fase. Na sexta fase, visto que o valor do ativo da empresa é inferior ao valor do passivo, observa-se a falência total ou falência técnica. Por fim, na sétima fase observa-se a falência, altura em que o processo de liquidação dos ativos começa para indemnizar os credores.

Perante as diferentes legislações em vigor em diferentes países, e baseando a interpretação do termo falência em conceitos legais, Altman, Iwanicz-Drozdowska, Laitinen e Suvas (2017), numa perspetiva internacional, constataram que a definição de falência não era homogénea mesmo entre países da mesma zona como os europeus. E com o propósito de evitar dúvidas, os autores classificaram as empresas falidas de acordo com o seu *status* na base de dados Orbis Europa (oferecida pela Bureau van Dijk), na qual falência é definida como a incapacidade explicita da empresa em pagar os seus credores, i.e., os ativos são vendidos para ressarcir os credores e, conseqüentemente, a empresa é extinta. Como será essa a base de dados que será usada no desenvolvimento deste trabalho, adotaremos a mesma classificação para o estudo que será desenvolvido, visto que a fonte de dados também é a Orbis Europa.

## 2.2 Estudos/Modelos de previsão de Falência

Ferramentas como os modelos de previsão de falências, que permitem estimar a situação económica e financeira da empresa, pelo menos, nos anos mais próximos, podem contribuir para diminuir a própria probabilidade de falência no sentido de auxiliar a empresa a tomar medidas preventivas eficientes. A falência da WorldCom Inc., uma das maiores fraudes no contexto empresarial da história com a falsificação de \$11 mil milhões nos resultados operacionais, é um exemplo claro da importância dessas ferramentas e a sua aplicação preventiva.

Nesse contexto, foram desenvolvidos ao longo do tempo vários modelos sobre a previsão de falência empresarial, tendo como referência os trabalhos pioneiros de Beaver (1966), Altman (1968) e Ohlson (1980), com a motivação comum de apurar a capacidade preditiva dos rácios financeiros para estimar a probabilidade de eventos de falência. Deste modo, a literatura que se seguiu pode ser classificada em: (1) Modelos Univariados, (2) Modelos Multivariados e (3) Modelos Multivariados de Probabilidade Condicionada.

Os modelos Univariados são aplicados analisando as variáveis escolhidas uma de cada vez, por via de um teste de classificação dicotómica que agrupa observações da amostra como falidas ou não falidas. Sendo uma análise caracterizada pela utilização de uma única variável para fins preditivos (Cook & Nelson, 1998).

Beaver (1966), inspirado pelo trabalho preliminar de autores como Fitz Patrick (1932), Winakor e Smith (1935) e Merwin (1942), recorreu à análise univariada para explicar a capacidade preditiva dos rácios financeiros para eventos como falência empresarial. O seu estudo consistiu em duas hipóteses: “1. Os rácios financeiros podem ser usados para prever falência. 2. Os referidos rácios não podem ser usados aleatoriamente, já que alguns podem ser mais exatos do que outros”, (conforme citado em Cook e Nelson, 1998, p.205-206). Ao analisar os rácios mais importantes para prever falência, i.e., os rácios mais significativos em agrupar as empresas em falidas e não falidas, o Beaver (1966) selecionou os que classificavam as empresas da amostra com menor percentagem de erro. Para tal, o autor usou uma amostra composta por 158 empresas dos EUA, escolhidas por um processo de *matching* por setor e dimensão, que operavam em 38 setores diferenciados, das quais 79 falidas e igualmente 79 empresas não falidas, durante o período de tempo entre 1954 e 1964.

Por sua vez, testou individualmente a capacidade de 30 dos rácios (subdivididos em 6 classes), com base na comparação dos valores médios anuais dos rácios financeiro para os 5 anos antes da falência. Com a informação dos rácios recolhida, o autor ordenou a mesma por ordem crescente e analisou-a com o objetivo de definir o *cutoff point* (i.e., o ponto que minimiza a percentagem de previsões incorretas) e classificar as empresas conforme definido *a priori*, empresas falidas e não falidas. Comparando os resultados com a classificação atual das empresas, determinou-se a percentagem de erro, i.e., quanto menor o erro maior o poder preditivo do modelo. Dos rácios testados, Beaver (1966) constatou uma grande diferença nas médias dos 2 grupos de empresas, tornando-se mais evidente a sua deterioração à medida que o ano da falência se aproximava.

O autor concluiu que, com uma capacidade preditiva de 87% um ano antes e 78% cinco anos antes da falência, o rácio  $\frac{Cash\ flow}{Total\ debt}$  demonstrou maior capacidade preditiva, apresentando uma percentagem de erro entre 13% e 22%, respetivamente, que são muito inferiores em relação ao esperado em modelos de previsão aleatória, onde a percentagem de erro esperada é de cerca de 50%. O rácio  $\frac{Net\ income}{Total\ assets}$  mostrou ser o 2º melhor como era expectável, visto que a sua correlação com o rácio  $\frac{Cash\ flow}{Total\ debt}$  é maior do que com qualquer outro rácio. O rácio  $\frac{Total\ Debt}{Total\ assets}$  é o terceiro com maior capacidade preditiva, seguido dos rácios de liquidez (*working capital to total assets*, *current ratio* e *no-credit interval*). Os resultados de Beaver (1966) evidenciam que os rácios são determinantes de eventos como falência, pelo menos até 5 anos antes da falência, reforçando as conclusões dos estudos que lhe serviram com base.

O estudo seminal de Beaver (1966) ofereceu um importante ponto de partida e, não obstante dos resultados obtidos terem sido satisfatórios, o carácter individualista da análise univariada levanta dúvidas e impõe limitações relacionadas com a compreensão dos resultados por considerarem os rácios isoladamente, ignorando qualquer correlação existente entre os mesmos. Dessa forma, o autor levantou a hipótese de que uma análise multivariada com a análise de vários rácios diferentes ao longo do tempo previsse melhor a situação de falência do que a análise aos rácios individuais.

Por sua vez Altman (1968), com um objetivo semelhante ao de Beaver (1966), investigou o tema recorrendo a uma metodologia diferente para ultrapassar as

dificuldades apontadas pelos modelos univariados, i.e., o autor optou pela Análise Discriminante Múltipla (ADM), que procura encontrar uma combinação linear de fatores, ou seja, rácios financeiros que melhor diferenciam as empresas falidas comparativamente às não falidas, através de uma função discriminante que o autor designou por  $Z^2$  ou valor *Z-Score*, e cuja principal vantagem consiste em considerar a evolução conjunta dos diferentes rácios, o que mitiga o enviesamento provocado pela possível manipulação das demonstrações financeiras pelos gestores.

De uma amostra cuidadosamente selecionada para garantir a homogeneidade relativamente à dimensão das empresas da amostra, foram analisadas 66 empresas cotadas do setor industrial, 33 falidas e 33 não falidas no período de 1946-1965. Altman (1968) considerou na sua análise um conjunto de 22 rácios, classificados em 5 categorias, abrangendo a Liquidez, rentabilidade, alavancagem, solvabilidade e atividade; desse conjunto foram apontados 5 rácios como sendo os mais eficientes na previsão de falência empresarial, sendo os rácios de rentabilidade ( $\frac{EBIT}{Ativo\ Total}$ ,  $\frac{Vendas}{Ativo\ Total}$  e  $\frac{Valor\ de\ Mercado\ do\ Capital\ Próprio}{Capital\ Alheio}$ ) os que mais contribuíram para a discriminação da função. A função discriminante estimada foi a seguinte:

$$Z = 0,012X_1 + 0,014X_2 + 0,033X_3 + 0,006X_4 + 0,999X_5$$

Onde:

$X_1 = Fundo\ de\ Maneio/Ativo\ Total$

$X_2 = Reservas\ e\ Resultados\ Transitados/Ativo\ Total$

$X_3 = EBIT/Ativo\ Total$

$X_4 = Valor\ de\ Mercado\ do\ Capital\ Próprio/ Capital\ Alheio$

$X_5 = Vendas/ Ativo\ Total$

O primeiro rácio reflete a dimensão e o nível de liquidez da empresa, sendo esse rácio apontado por Altman (1968) como o melhor indicador para a discriminação das empresas na análise multivariada assim como na análise univariada, o que é consistente com o estudo de Merwin (1942). O segundo rácio que apresenta maior significância estatística, reflete a capacidade/desempenho financeiro da empresa e considera (implicitamente) a idade da empresa. Segundo o autor uma empresa mais

---

<sup>2</sup> Função discriminante por Altman (1968):  $Z = V_1X_1 + V_2X_2 + \dots + V_nX_n$ , onde  $V_1, V_2, \dots, V_n$  = Coeficiente Discriminante e  $X_1, X_2, \dots, X_n$  = Variáveis Independentes

jovem possivelmente apresentará baixos resultados transitados comparativamente a uma empresa mais antiga porque não teve tempo suficiente para aumentar os seus lucros acumulados. O terceiro rácio é o indicador da rendibilidade do ativo da empresa, não obstante de quaisquer impostos ou responsabilidades financeiras. O quarto rácio mede a depreciação dos ativos da empresa antes que ocorra uma eventual falência. Por fim, o quinto rácio designado por rácio de rotação do ativo reflete a capacidade de geração de vendas pelos ativos utilizados. Independentemente deste último rácio ter apresentado um nível de significância inferior na Análise Univariada, foi uma descoberta surpreendente para o autor ao aparecer como o segundo maior contributo para discriminar as empresas falidas das não falidas.

Para um melhor entendimento do *Z-score* e classificação das empresas saudáveis ou falidas, as empresas classificadas incorretamente foram analisadas tendo em conta diferentes valores de *Z-score*, ou seja diferentes pontos de *cutoff* (o *score* que minimiza o número de classificações incorretas), constatando-se que todas as empresas saudáveis apresentam um *Z-score* superior a 2,99, enquanto as empresas falidas apresentam um *Z-score* inferior a 1,81. Foram também avaliados diferentes intervalos entre o *score* de 1,81 e 2,99 até encontrar o ponto médio do intervalo de 2,675 que é o *cutoff point*. A zona entre 1,81 e 2,99 foi definida como a “zona cinzenta ou zona de incerteza” devido à sensibilidade na classificação de erros do tipo I e II dado que esses erros (falsos positivos e falsos negativos) são observados nessa zona.

Como uma percentagem de acerto de 95% da amostra um ano antes da falência acontecer, o modelo de Altman (1968) demonstrou um elevado nível de eficiência. No entanto, ao alargar os prazos para 2, 3, 4 e 5 anos antes da falência, observou-se uma redução da capacidade de previsão do modelo para 72%, 48%, 29% e 36%, respetivamente. Esta evidência contribui para a conclusão que a deterioração dos rácios é cada vez mais evidente com a aproximação da data da falência, conforme foi verificado na análise univariada (p.e., Beaver 1966 e Merwin 1942). Além disso, verificou-se ainda que as mudanças mais preocupantes na maioria dos rácios analisados ocorriam entre o terceiro e segundo ano antes da falência. Mais tarde, Altman, Haldeman e Narayanan (1977), para testarem a atualidade do *Z-score*,

desenvolveram o modelo ZETA<sup>3</sup>. Para um universo amostral de 111 empresas, 53 falidas no período entre 1969 e 1975, foram utilizadas 7 variáveis independentes e analisado o efeito da dimensão nas mudanças estruturais das empresas. Mantendo o *cutoff* inicial de 2,675, os autores constataram que o modelo continuou a apresentar um elevado poder preditivo, apesar de registrar uma pequena queda em relação ao *Z-score* ao apresentar uma capacidade de previsão de 90% para um ano antes da falência, e cerca de 70% para 5 anos antes da falência.

Pese embora os resultados apresentados serem satisfatórios, o *Z-score* de Altman (1968) era restrito a empresas cotadas. Assim, Altman (1983) ao propor um modelo mais abrangente, substituiu a variável que representa o valor de mercado da empresa ( $X_4$ ) pelo seu valor contabilístico. A partir das estimativas obtidas, verificou-se que a capacidade preditiva do modelo mantém-se em padrões semelhantes ao anterior, embora com um aumento na incerteza da precisão do modelo devido a uma pequena redução da sensibilidade de 94% para 91% e o aumento da amplitude da “zona cinzenta” visto que o seu limite inferior passou a ser 1,23 em vez de 1,81. Tal como no estudo publicado em 1968, neste novo estudo, o autor constatou que o rácio que demonstrou maior poder preditivo foi o  $X_2 = \frac{\text{Reservas e Resultados Transitados}}{\text{Ativo Total}}$

Inspirado pelos resultados satisfatórios do *Z-score* de Altman, Deakin (1972) obteve resultados semelhantes ao recorrer a uma metodologia multivariada e ao testar os 14 dos 30 rácios de Beaver (1966), para uma amostra composta por 97 empresas saudáveis e 97 empresas falidas durante 3 anos, desenvolveu um modelo com capacidade de previsão de falência de 94% um ano antes do evento.

Na sequência da metodologia Multivariada, Blum (1974), desenvolveu o seu modelo de previsão de falência para um universo amostral de 115 empresas falidas e 115 empresas saudáveis analisadas entre 1954 e 1968, emparelhando cada empresa falida a uma empresa não falida com a mesma dimensão e o mesmo setor de atividade, constatou que a capacidade preditiva do seu modelo diminuiu à medida que aumenta o número de anos antes do evento de falência. Resultados esses que se apresentaram semelhantes aos estudos empíricos anteriores. E tal como Beaver (1966), Blum concluiu na sua investigação que o rácio que apresentou maior poder preditivo na

---

<sup>3</sup> As 7 variáveis do modelo ZETA proposto por Altman et al. (1977)  $X_1 = \frac{\text{Resultados Operacionais}}{\text{Ativo Total}}$ ;  $X_2 = \frac{\text{Desvio padrão da rentabilidade do ativo, num período entre 5 e 10 anos}}{\text{Reservas e Resultados Transitados}}$ ;  $X_3 = \frac{\text{Resultados Operacionais}}{\text{Custos Financeiros}}$ ;  $X_4 = \frac{\text{Ativo de Curto Prazo}}{\text{Passivo de Curto Prazo}}$ ;  $X_5 = \frac{\text{Capital Próprio}}{\text{Capital Total}}$ , e  $X_7 = \text{Log}(\text{Ativo total})$ .

explicação da situação de falência foi o  $\frac{Cash\ Flow}{Divida\ total}$ .

Altman et al. (1979) com um objetivo idêntico e analisando uma amostra de empresas do mercado Brasileiro, utilizaram a base do *Z-score* de Altman (1968) para desenvolver um modelo para prever a ocorrência de eventuais problemas financeiros, recolheram dados para uma amostra composta por 58 empresas, sendo que 23 dessas empresas apresentavam indícios de problemas financeiros. Os resultados permitiram concluir que para 1 ano e 3 anos antes da ocorrência do evento, o modelo classificava corretamente as empresas com uma precisão de 88% e 83%, respetivamente.

O *Z-score* de Altman (1968) é, ainda hoje, um dos modelos mais utilizados para prever a situação de falência empresarial. No entanto, o modelo apresenta algumas limitações, sendo alguns dos pressupostos inerentes à metodologia ADM a maior das críticas. A título de exemplo, Ohlson (1980) apontou o factos como : i) O modelo *Z-score* assumir a distribuição normal das variáveis e a igualdade das matrizes de variância e covariância entre os dois grupos de empresas ( falida e saudáveis); ii) O resultado, i.e., o *score* obtido na análise discriminante não é de fácil interpretação, visto a análise faz apenas a discriminação das empresas entre os dois grupos e não ilustra a probabilidade de uma empresa entrar em falência; iii) E ainda o facto de ao recorre-se a técnica de matching, as empresas falidas e saudáveis são comparadas por critérios considerados pouco arbitrários como o setor da atividade e a dimensão .

Consciente dessa e de outras limitações do ADM, Ohlson (1980) utilizou um modelo multivariado de probabilidade condicionada, mais concretamente, recorrendo à análise logística, para desenvolver o seu modelo de previsão da falência. A análise Logit caracteriza-se por relaxar alguns pressupostos da ADM, tais como o pressuposto da normalidade das variáveis, o que possibilita a análise com amostras desproporcionais. Esta metodologia parece adequar-se bem a estudos de previsão de falência onde a probabilidade de falência vs não falência é uma variável binária; os grupos de empresas falidas e não falidas assumem o valor de 1 ou zero, não se sobrepõem e são diferenciáveis. E sendo um modelo probabilístico a questão fundamental da estimativa resume-se à probabilidade de uma empresa cair numa situação de falência num determinado espaço de tempo.

A partir de uma amostra de 2.163 empresas (2.058 empresas saudáveis e 105 empresas falidas) entre 1970 a 1976, tendo em conta alguns fatores como a exclusão

de empresas não cotadas, empresas públicas, empresas financeiras, companhias de transporte, entre outras empresas pertencentes a setores muito regulados (p.e., setor da energia), Ohlson (1980) propôs o seu modelo, baseado no recurso a um modelo *logit* de previsão de falência, utilizando 9<sup>4</sup> variáveis, das quais 2 variáveis binárias e 7 variáveis financeiras. Com base nessas variáveis, o autor analisou a probabilidade de falência 1 antes, 2 anos antes e 3 anos antes, e não obstante do resultado do modelo desenvolvido para 1 ano antes da falência ter sido inferior, a capacidade preditiva para 1 e 2 anos antes do evento de falência foram superiores a 90%, atingindo assim uma capacidade preditiva superior ao modelo de Altman(1968).

Com base nos resultados obtidos, Ohlson (1980) constatou a existência de 4 indicadores estatisticamente significativos para explicar a situação de falência, sendo estas: a dimensão da empresa (SIZE), a estrutura financeira (TLTA), a performance (NITA) e a liquidez geral (CLCA), para um e dois anos antes da falência.

Embora o modelo desenvolvido por Ohlson (1980) tenha apresentado uma capacidade preditiva inferior em relação a estudos anteriores que recorreram à metodologia ADM, foi uma “lufada de ar fresco” na metodologia aplicada para previsão da falência das empresas, inspirando autores como Zmijweski (1984), Zavgren (1985) e Platt (1990), que recorreram a mesma metodologia nos seus trabalhos. No entanto, Zmijweski (1984) recorreu ao modelo de probabilidade condicionada Probit, que difere do *logit* por utilizar uma função de probabilidade acumulada e os preditores do modelo seguem uma distribuição normal padronizada. De uma amostra de 800 empresas saudáveis e 40 falidas, para o período entre 1972 a 1978, o autor utilizou apenas o rácio de alavancagem, o rácio de liquidez e o rácio de rentabilidade do ativo evidenciando que o desempenho do seu modelo *probit* era similar aos modelos já desenvolvidos. Posteriormente aos trabalhos de Beaver (1966), Altman (1968) e Ohlson (1980), vários outros modelos de previsão de falência de empresas foram desenvolvidos e demonstraram um elevado poder preditivo. No entanto, num estudo realizado por Bellovary, Giacomino e Akers (2007), onde analisaram a modelização da falência de empresas, os autores verificaram que

---

<sup>4</sup> As variáveis selecionadas por Ohlson(1980):  $SIZE = \log \frac{Total\ Assets}{GDP\ price\ Index}$  (variável usada como proxy da dimensão da empresa)  $TLTA = \frac{Total\ Liabilities}{Total\ Assets}$ ;  $WCTA = \frac{Working\ Capital}{Total\ Assets}$ ;  $CLCA = \frac{Current\ Liabilities}{Current\ Assets}$ ;  $OENEG = dummy$  que apresenta valor 1 se o passivo for superior ao activo, 0 caso contrário;  $NITA = \frac{Net\ Income}{Total\ Assets}$ ;  $FUTL = \frac{Cash\ Flow}{Total\ Liabilities}$ ;  $INTWO = dummy$  que apresenta valor 1 se o resultado líquido dos últimos 2 anos for negativo, 0 caso contrário;  $CHIN = \frac{Net\ Income_t - Net\ Income_{t-1}}{|Net\ Income_t| + |Net\ Income_{t-1}|_{total}}$ .  
 Modelo proposto por Ohlson(1980):  $O-Score = \alpha + \beta_1 SIZE + \beta_2 TLTA + \beta_3 WCTA + \beta_4 CLCA + \beta_5 OENEG + \beta_6 NITA + \beta_7 FUTL + \beta_8 INTWO + \beta_9 CHIN$



dos mais de 150 modelos desenvolvidos no decorrer das últimas décadas, muitos destes foram aplicados a grandes empresas, pressupondo a grande quantidade e a boa qualidade de dados disponíveis, defendida na generalidade.

Contudo, temos assistido a um crescente de trabalhos académicos que se focam no desenvolvimento de modelos aplicados ao contexto e realidade nacional, como é o caso de Sera (2020), Pereira (2021) e Silva (2021) que desenvolveram modelos de previsão de insolvência para PME portuguesas.

## **2.3 Definição e Estudos de PME**

### **2.3.1 Conceito de PME**

Para melhor compreensão do tópico em estudo, é importante clarificar o conceito de PME de modo a evitar incoerências dado que este conceito varia de país para país. Mas não obstante dessa variabilidade, algumas das características mais utilizadas para definir PME são volume de negócios anual, o valor total dos ativos, o número de funcionários, proveitos anuais, capital investido, forma jurídica da empresa, tipo de indústria, entre outras (Altman et al. (2010)).

Atualmente na UE, as PME são definidas com base na recomendação 2003/361/CE de 6 de maio, , segundo o qual são consideradas PME todas as empresas que correspondem aos seguintes requisitos: número de empregados inferior a 250, volume anual de negócios inferior a 50 milhões de EUR ou saldo do balanço (i.e., total do ativo) inferior a 43 milhões de euros. E tendo em conta o universo geográfico em estudo – Portugal e Espanha, adotaremos a definição em vigor na UE. A Tabela 1 apresenta o detalhe da classificação das PME.

**Tabela 1 – Definição de PME**

<b>Categoria da Empresa</b>	<b>Nº de empregados</b>	<b>Volume de Negócios</b>		<b>Balanco Total</b>
Médias empresas	< 250	≤ 50 milhões de euros	OU	≤ 43 milhões de euros
Pequenas empresas	< 50	≤ 10 milhões de euros	OU	≤ 10 milhões de euros
Micro empresas	< 10	≤ 2 milhões de euros	OU	≤ 2 milhões de euros

**Nota:** Fonte: Adaptado do Annual Report on European SME 2018/2019

### 2.3.2 Estudos de PME

A existência de vários indicadores macroeconómicos que enfatizam a importância das PME para o desenvolvimento da economia, despertou o interesse de muitos investigadores em prever possíveis eventos de falência para esse segmento de empresas, sobretudo na UE depois da crise da dívida soberana entre 2010/2013, sendo que em alguns países, a implementação de medidas de austeridade causou um impacto negativo no crescimento da maioria das PME, além de ter conduzido a um número elevado de falências.

Desde os primeiros estudos sobre a previsão de falência que as PME têm vindo a ser ignoradas. Embora a evidência apontasse que a probabilidade de falência será maior nas pequenas empresas (Beaver,1966), as amostras analisadas eram sempre referentes a grandes empresas. Edmister (1972) foi pioneiro no desenvolvimento de um modelo de previsão de falência específico para PME. Neste contexto, o autor utilizou a metodologia ADM e testou 19 variáveis, para o período entre 1954 e 1969, numa amostra de 562 e outra de 42 empresas para um horizonte temporal de 1 e 3 anos consecutivamente, conseguindo desenvolver um modelo que demonstrou um elevado poder preditivo na deteção do evento de falência.

No entanto, dada a crescente importância das PME na economia, e inspirados na investigação de Edmister (1972), Altman e Sabato, em 2007, construíram um modelo de previsão de falência com base em informação um ano antes da situação de falência, destinado as PME dos EUA e usando a metodologia de regressão logística. No seu estudo, os autores a partir de uma amostra 2010 empresas saudáveis e 120

empresas falidas, com um volume de negócios inferior a 65 milhões de dólares<sup>5</sup>, no período 1994-2002, testaram 16 rácios financeiros, dos quais 5 (liquidez, a alavancagem, a atividade, a rentabilidade e o rácio de cobertura) apresentaram-se como os mais relevantes na previsão de falência das PME. O modelo de Altman e Sabato (2007) para PME demonstrou uma capacidade preditiva 30% superior quando comparado com um modelo genérico para grandes empresas, i.e., demonstraram que a eficiência de um modelo logístico específico para PME é consideravelmente superior a de um modelo que usa a ADM para grandes empresas. Adicionalmente, Altman e Sabato (2007) efetuaram uma transformação logarítmica tendo em conta a particularidade de cada variável do conjunto das que revelaram elevado poder preditivo com o objetivo de aumentar a eficácia do modelo e observaram que a capacidade de previsão do modelo aumentava de 75% para 87%. Para além dos resultados satisfatórios obtidos, os autores notaram que a separação entre PME e grandes empresas é ainda mais relevante para o setor bancário, deixando assim evidente a importância do desenvolvimento de modelos específicos para avaliar a probabilidade de falência para empresas de menor dimensão.

Apesar do modelo desenvolvido ter se mostrado eficiente, em 2010, Altman, Sabato e Wilson, a fim de melhorar o poder preditivo do mesmo, estenderam-no incluindo variáveis não financeiras que espelham as particularidades da empresa e aspetos importantes sobre a situação da mesma; os autores testaram esse novo modelo usando uma amostra de PME do Reino Unido. Assim, os autores selecionaram 5.749.188 empresas (66.833 falidas e 5.682.355 saudáveis), para o período entre 2000-2007 e tendo como base o modelo logit desenvolvido em 2007, demonstraram que a inclusão de variáveis não financeiras aumenta a precisão do modelo de previsão de falência até 13% um ano antes do evento, o que reforça a solidez e o poder de previsão significativo deste modelo específico para prever falência de PME. Nessa sequência, destacando a importância da inclusão de variáveis qualitativas para além das quantitativas na modelização de falência, Blanco, Irimia e Oliver (2012) recorreram também à metodologia logística e construíram o seu modelo de previsão de falência para PME não cotadas do Reino Unido um ano antes do evento de falência ocorrer. Os autores analisaram uma amostra composta por 38.570 PME, para o

---

<sup>5</sup> Segundo a recomendação da UE, uma empresa é considerada média caso o volume de negócios seja inferior a 50 milhões de euros ou 65 milhões de dólares.

período temporal entre 1999 e 2008. Considerando variáveis financeiras, não financeiras e macroeconómicas, os resultados obtidos nesta investigação demonstraram mais uma vez que a junção de variáveis financeiras e não financeiras melhora o poder preditivo na modelização da previsão de falência para PME.

Visto que os primeiros estudos não continham variáveis qualitativas, o surgimento de estudos como Altman *et al.* (2010) e Blanco *et al.* (2012), que demonstram evidência que a inclusão de variáveis qualitativas aumenta a capacidade preditiva dos modelos de previsão de falência, abriram caminho à procura por informação não financeira que ajude a explicar a situação de falência. Porém, Dănilă (2012) argumenta que essa melhoria dependia principalmente da qualidade e do tratamento dos dados qualitativos.

Tendo em conta um mercado cada vez mais globalizado, Gupta, Wilson, Gregoriou e Healy (2014), motivados pela controvérsia existente na literatura sobre o desempenho das PME internacionais, utilizaram a análise logística para desenvolver um modelo de previsão de falência das PME no Reino Unido, com a pretensão de incluir variáveis que refletissem o impacto da internacionalização nesse segmento de empresas. Classificando as empresas com o estatuto de internacional se reportarem proveitos operacionais relativos à atividade exportadora (e doméstica, caso contrário), os autores consideraram uma amostra de 342.711 PME domésticas (com 8.553 falidas) e 344.205 PME internacionais (com 9.114 falidas). As variáveis que capturam o efeito da exportação sobre a probabilidade de falência das PME internacionais não foram consideradas no modelo final dado não apresentarem significância estatística; os autores demonstraram deste modo que o conjunto de fatores que afeta a probabilidade de falência é o mesmo para os dois grupos, entrando em contradição com estudos como o de Arslan e Karan (2009), que sugeriam que ao desenvolver modelos de previsão de falência para PME devem ser consideradas, separadamente, as empresas domésticas e internacionais.

Mais recentemente, em 2017, Altman, Iwanicz-Drozdowska, Laitinen & Suvas, também para um contexto internacional, analisaram o *Z-score* de Altman (1983) utilizando uma amostra de 31 países europeus e 3 não europeus. A amostra é composta principalmente por empresas privadas (financeiras e não financeiras) de todos os setores de atividade constituída por 2.602.563 empresas saudáveis e 38.215 empresas falidas entre 2002 e 2010. Apresentando uma capacidade preditiva

superior a 90% com a inclusão das variáveis macroeconómicas que diferenciavam os países em análise, o modelo *Z-score* de Altman (1983) demonstra mais uma a sua predominância face a outros ao refletir capacidade preditiva numa amostra *cross-country*. Os resultados desse estudo permitem concluir que tal como acontece na análise de previsão de falência de PME em contexto internacional, não há necessidade de se analisar as grandes com um modelo ajustado para esse contexto específico, complementando a conclusão de Gupta *et al*, (2014).

A ideia de que os modelos de previsão de falência são ferramentas úteis que os agentes económicos podem aplicar no contexto real, fez com que eles ganhassem cada vez mais destaque, enriquecendo a literatura ao explorar metodologias estatísticas alternativas como é o caso da *Support Vector Machine (SMV)*<sup>6</sup>, das redes neuronais artificiais (RNA) e do modelo logístico de múltiplos períodos (*discrete hazard*), para melhorar a precisão dos modelos de previsão de falência.

Neste contexto, de metodologias alternativas para a previsão de falência, Lacerda e Moro (2008), usando três metodologias de estimação: regressão logística, análise discriminante e *Support Vector Machine*, estudaram a probabilidade de falência das empresas portuguesas. Os autores constataram que comparando as três técnicas utilizadas, a SVM revela ser superior, o que pode ser explicado pela contemporaneidade tecnológica do modelo, que permite utilizar um conjunto mais vasto de indicadores capazes de prever a falência das PME. Outra razão apontada pelos autores para a superioridade da SVM é a sua flexibilidade para capturar a dependência não monótona dos preditores da probabilidade de falência. Dos 47 rácios testados, os que revelaram maior capacidade preditiva (analisados na perspetiva dos 3 métodos utilizados) são os que caracterizam a estrutura de custos, a liquidez, a atividade e a alavancagem, sendo que foram ainda testados outros indicadores financeiros, que se mostraram preditores eficientes, tais como o tempo médio de emprego dos funcionários e o aumento do número de relacionamentos bancários que é um indicador que nunca foi utilizado nesse tipo de estudos.

Aplicando a metodologia das redes neuronais artificiais, Du Jardin (2010) construiu um modelo de previsão de falência aplicando diferentes métodos de seleção de variáveis, com o propósito de analisar a sua influência na precisão do modelo de

---

<sup>6</sup> Segundo (Lacerda & Moro, 2008), *Support Vector Machine (SVM)* é uma metodologia de classificação não linear, que se baseia na separação de duas classes de observações: empresas falidas e empresas não falidas.

falência. De acordo com o autor, um modelo para prever a situação de falência baseado em redes neurais, quando construído com métodos de seleção de variáveis apropriados em vez utilizar métodos frequentemente utilizados na literatura financeira, tem uma performance superior quando comparado aos modelos ADM e logístico.

Mais recentemente, e consciente dos grandes desafios que as economias em desenvolvimento enfrentam, Charalambakis e Garrett (2019) construíram um modelo de previsão de falência para PME Gregas no período entre 2003-2011, usando o modelo logístico múltiplo (*discrete hazard*), por ser um modelo que captura o impacto das mudanças da situação da empresa ao longo do tempo. Ao analisar dados que incluem a crise da dívida soberana que a Grécia sofreu entre 2009 a 2011, os autores desenvolveram um modelo que classifica corretamente 88% das empresas que faliram durante essa crise, assim como a probabilidade de um empresa falir num cenário que exclui a crise, demonstrando evidência que as falências que aconteceram durante a crise poderiam ser previstas através da análise de dados pré-crise.

Todos estes estudos apontam que a metodologia mais adequada para se utilizar em modelos de previsão de falência parece estar longe de estar concluída.

## **2.4 Objetivos do Estudo**

Considerado um tema de grande importância, particularmente nos dias de hoje, com a crise macroeconómica que o mundo vive devido à pandemia global causada pela Covid-19 e pelas medidas adotadas para a contenção da transmissão desse vírus, esses eventos de falência são inevitáveis e acredita-se que podem vir a gerar uma recessão ainda pior do que a vivida anteriormente.

Nesse contexto, tendo em conta o papel primordial das PME para a recuperação da economia europeia e do mundo, o objetivo principal deste trabalho consiste em desenvolver um modelo de previsão de falência específico para PME na Península Ibérica, aplicando como metodologia base a regressão logística de acordo com Ohlson (1980), dado que do ponto de vista estatístico esta metodologia parece adequar-se bem às características do problema de predição da situação de falência, onde a variável dependente é binária (empresas falidas vs empresas não falidas) (Altman & Sabato, 2007).

A escolha pelo conceito de falência em detrimento de insolvência deve-se ao caráter internacional deste trabalho, que inclui os mercados de Espanha e Portugal, onde existem diferenças em relação à lei que regula o processo de insolvência/falência empresarial. Deste modo, o conceito falência é mais abrangente e melhor percecionado como uma situação que pode, não só, conduzir a um processo jurídico, mas sobretudo a constrangimentos financeiros severos.

Nesse sentido, para a construção do modelo semelhante ao proposto por Ohlson (1980) e ajustado à realidade que se pretende estudar, serão construídas variáveis na forma de rácios económico-financeiros de acordo com a literatura anterior (p.e, Jackendoff, 1962; Beaver, 1966; Altman, 1968; Ohlson, 1980; Altman e Sabato, 2007; Lacerda e Moro, 2008; Altman et al., 2010) e analisadas a fim de serem selecionadas as variáveis que apresentem maior capacidade explicativa perante a situação de falência de PME na Península Ibérica. Sendo que para tal, recorrer-se-á a diferentes metodologias com a Análise Univariada, Análise Discriminante Multivariada (ADM), Análise de Regressão Logística e o método Propensity Score Matching (PSM) proposto por Rosenbaum e Rubin (1983).

Por fim, a eficácia do modelo proposto será comparada ao modelo *Z-Score* (específico para grandes empresas) de Altman (1968) na expectativa de que a percentagem de acerto do modelo produzido seja superior ou esteja entre 70% e 99%, como desejável na previsão da falência das empresas (Laffarga, 1999).

### 3. Metodologia

De acordo com o objetivo primário deste trabalho, será definida a variável dependente como binária (0 se a empresa estiver classificada como falida e 1 se a empresa estiver classificada com não falida) e serão descritas as variáveis independentes definidas *a priori*. As variáveis independentes serão analisadas de forma univariada para se avaliarem as diferenças entre os grupos de empresas falidas e empresas saudáveis e entre os países da Península Ibérica (Espanha *Versus* Portugal). De seguida, será conduzida a análise multivariada. Serão aplicadas as metodologias usadas por Altman (1968) e Ohlson (1980), que apesar das suas diferenças, têm por objetivo testar diferenças entre grupos onde um dos grupos sofreu um dado tratamento ou evento, por exemplo, falência, e o outro grupo não sofreu esse tratamento, sendo considerado o grupo de controlo.

Ainda, para garantir a consistência e a segurança dos resultados da análise, será aplicada a técnica de Propensity Score Matching (PSM) proposta por Rosenbaum and Rubin (1983), que procura emparelhar empresas com características comuns, associando uma empresa falida e uma empresa não falida, de modo a reduzir o enviesamento causado pela desproporcionalidade da amostra de empresas falidas *versus* saudáveis.

#### 3.1 Análise Discriminante Múltipla

A Análise Discriminante Múltipla (ADM) é uma metodologia estatística cujo objetivo é classificar qualquer observação em um ou mais grupos *a priori*, dependendo das características individuais de cada observação, sendo neste caso o grupo de empresas falidas vs grupo de empresas não falidas. Este método será utilizado como análise preliminar que irá complementar a análise logística, que será a metodologia principal.

O primeiro modelo de previsão de falência baseado na ADM foi o famoso *Z-Score*, proposto por Altman em 1968, com objetivo de identificar fatores determinantes da situação de falência utilizando para esse efeito um conjunto de variáveis económico financeiras em simultâneo. A ADM envolve a definição de uma função linear apresentada da seguinte forma:



$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

Onde  $Y$  é o *score* da função discriminante,  $\beta_j$  são os coeficientes da função e  $X_n$  as variáveis independentes.

A análise discriminante estabelece diferentes características que distingam os grupos que estão a ser analisados e, de seguida, os coeficientes discriminantes são multiplicados pelas respetivas variáveis independentes conforme a função discriminante, obtendo-se dessa forma um índice ou *score*, que é utilizado para classificar se uma empresa se encontra ou não em situação de falência de acordo com os grupos já definidos (ver, p.e.: SAGE, 2015, pp. 590-592).

A aplicação da ADM, deve respeitar alguns pressupostos, tais como: i) distribuição normal das variáveis, ii) a igualdade das matrizes de variância e covariância, iii) a existência de diferenças significativas entre os grupos e iv) a remoção de *outliers* (Ohlson, 1980). No entanto, quanto ao pressuposto de que as variáveis económicas e financeiras seguem uma distribuição normal, raramente isso é observado, pelo que essa hipótese é rejeitada na maioria dos casos (Ohlson, 1980). Já no que respeita à igualdade das matrizes de variância e covariância, essa hipótese não é rejeitada, visto ser uma consequência da linearidade da função. É importante a escolha de variáveis que apresentem diferenças significativas nas médias entre grupos (falidas e não falidas) e essa seleção prévia pode ser realizada usando diferentes técnicas estatísticas, como o recurso à análise univariada na medida em que esse processo influenciará os resultados. Quanto ao último pressuposto da existência de *outliers*, este fenómeno deve ser tratado antes da aplicação da metodologia dado que a sua presença pode ter um impacto negativo na aplicação da ADM; a observação de valores médios e de dispersão deturpados devido aos dados anormais podem enviesar os resultados obtidos a partir de uma qualquer amostra.

Não obstante dos pressupostos rígidos impostos pela ADM, esta metodologia aplicada à análise de previsão de falência apresenta outras limitações, de onde se destaca o facto de ser considerada uma análise baseada exclusivamente em informação contabilística e financeira por não incluir dados não financeiros (como, por exemplo, as competências do gestor, formação dos trabalhadores, entre outras), e ainda pelo facto de que para ser utilizada é necessário saber o momento exato que antecede a falência.

Considerado ser o método multivariado mais usado desde os anos 60 para prever falência, a ADM é uma técnica facilmente aplicável, que apresenta resultados confiáveis e vantagens, como o facto de ser uma análise que permite a combinação de vários rácios financeiros para uma melhor perceção sobre o desempenho financeiro das empresas, do ponto de vista global, sendo mais eficiente na sua função do que a análise univariada (Altman, 1968). Ainda segundo Morgado (1998), a ADM apresenta a vantagem de contribuir para a mitigação do efeito que pode surgir devido à provável manipulação de dados contabilísticos por parte dos gestores.

### 3.2 Modelos Logísticos

A análise logística é outro método muito utilizado no desenvolvimento de modelos de previsão de falência, distinguindo-se como uma técnica de probabilidade condicional. A variável dependente ( $y$ ) é dicotómica, vulgarmente designada por *dummy*, que identifica as empresas como “saudável” ou “falida”, ou seja, é uma variável binária que pode assumir o valor de  $Y=1$  no caso de se observar um evento de falência, ou  $Y=0$  caso contrário.

Ohlson (1980), foi o primeiro a propor a metodologia de regressão logística para desenvolver um modelo de previsão de falência, uma vez que relaxa alguns dos pressupostos exigidos pela ADM em relação à distribuição normal das variáveis, apresenta vantagens como a possível inclusão de informação de carácter qualitativo e a utilização de uma amostra desproporcional entre o número de observações de empresas falidas e não falidas.

O modelo logístico é representado pela seguinte equação:

$$Prob(Y_i = 1) = \beta'Z_{i,t-1} + \lambda_k + \gamma_t + \varepsilon_{it} \quad (2)$$

Assim, o modelo logístico é estimado pelo método da máxima verosimilhança e calcula a probabilidade de ocorrência de um evento de falência ( $Y_i = 1$ ) em função de um conjunto de  $n$  variáveis independentes, desfasadas no tempo relativamente ao momento de ocorrência da falência ( $\beta'Z_{i,t-1}$ ), sendo que, os parâmetros  $\lambda_k$  e  $\gamma_t$  são considerados para controlarem a heterogeneidade não observada em relação à indústria e os anos incluídos no painel respetivamente.  $\varepsilon_{it}$  é o erro da regressão.

A probabilidade acumulada da função é calculada pela seguinte expressão:

$$Prob(Y_i = 1) = \frac{1}{1+e^{-x_i\beta}} \quad (3)$$

Além do logit, outro modelo de probabilidade condicional muito utilizado, cuja variável dependente é também uma *dummy*, é o Probit que também é estimado através do método da máxima verosimilhança e difere do modelo logit somente pelo facto da sua distribuição ser normal padronizada e a distribuição do modelo *logit* ser lognormal, o que implica que os coeficientes tenham que ser transformados em probabilidades marginais para serem interpretados (Zmijweski, 1984).

### 3.3 Propensity Score Matching

Sabendo que existe uma desproporcionalidade quanto ao número de observações de empresas falidas *versus* empresas não falidas, sentiu-se a necessidade de se aplicar um método complementar à análise principal que nos permitirá garantir a consistência e a segurança dos resultados da análise. Assim, o método complementar que será aplicado é a técnica de *Propensity Score Matching* (PSM) proposta por Rosenbaum and Rubin (1983), que será utilizado com o objetivo de ultrapassar o problema de enviesamento da amostra.

Para a aplicação da referida metodologia, serão utilizadas as variáveis definidas *a priori* e controlando para o ano e indústria. Este método estatístico procura emparelhar empresas com características comuns antes da ocorrência da falência, associando uma empresa falida e uma empresa não falida. Para tal é preciso criar um grupo de controlo, constituído pela amostra de empresas saudáveis e a um grupo de tratamento composto pela amostra de empresas falidas, fazendo corresponder cada observação do grupo de tratamento a uma observação do grupo de controlo.

Assim, numa primeira fase estima-se a probabilidade de uma empresa entrar numa situação de falência, selecionando um conjunto de variáveis explicativas da situação de falência denominadas *covariates* e, dessa forma, a estimação do *Propensity Score* é feita por meio da análise de um modelo probit. De seguida, aplica-se o *matching* – cada observação do grupo de empresas falidas será comparada a uma observação semelhante do grupo de empresas saudáveis, com auxílio do método

*nearest neighbour e with replacement*, em termos médios e durante o período anterior ao período da ocorrência da falência. É de salientar ainda que *with replacement* significa que existe a opção de uma empresa saudável ser usada como match para uma só empresa falida ou para várias empresas falidas, respetivamente.

A utilização da metodologia *Propensity Score Matching* além de equilibrar a amostra das empresas saudáveis face à amostra de empresas falidas, ajuda também a evitar que a eficácia do modelo a estimar não seja afetada pela referida desproporcionalidade da amostra.

## 4. Descrição da Amostra e das Variáveis

### 4.1 Descrição da Amostra

Considerando o objetivo deste trabalho empírico de desenvolver um modelo de previsão de falência para as PME da Península Ibérica (classificadas segundo recomendação da Comissão Europeia, de 6 de maio de 2003, relativa à definição de micro, pequenas e médias empresas), a análise será conduzida tendo em conta uma amostra composta por informação financeira referente ao segmento de empresas em questão e será recolhida da base de dados ORBIS Europe da Moody's Analytics, cobrindo o período entre 2012 e 2020, visto ser o horizonte temporal disponível na referida base de dados.

Neste estudo serão classificadas como falidas as empresas que, de acordo com a base de dados da ORBIS, apresentam estatus *Bankruptcy* (refere-se a empresas que estão numa fase do processo de falência em que o ativos da empresa são vendidos para pagar os credores, resultando na extinção da mesma), *Dissolved (bankruptcy)* (refere-se a empresas extintas), *Active (insolvency proceedings)* (nesta situação, a empresa ainda tem atividade e procura reembolsar os seus credores, embora apresente indícios de dificuldades financeira e o processo de falência esteja a decorrer no tribunal, sendo que a empresa poderá futuramente voltar à normalidade, ser reestruturada ou liquidada); e as empresas que apresentam status *Active* (refere-se ao grupo de empresas saudáveis) serão classificadas como não falidas, garantindo que os 2 grupos possuem características semelhantes.

Para a obtenção da amostra foram aplicados critérios como:

- a) Forma legal: Foram selecionadas empresas privadas e não cotadas;
- b) Indústria: Foram excluídas da análise atividades altamente reguladas, como o caso da distribuição de energia ou água (CAE 35 e 36) e o setor financeiro (CAE 64,65,66 e 68);
- c) Dimensão: Foram excluídas Médias, Pequenas e Micro Empresas com o número de trabalhadores inferior a 2.

Do conjunto de empresas previamente recolhidas, que cumpriam com os critérios acima mencionados, foram ainda eliminadas todas as observações que

apresentavam valor do Ativo Total igual ou inferior a zero, Volume de Negócios igual ou inferior a zero, e valores do Ativo Corrente, Passivo Corrente e Passivo não circulante inferior a zero.

No final, obteve-se uma amostra constituída por 16.536 PME (correspondente a 146.209 observações) segmentada da seguinte forma: i) grupo das empresas saudáveis 16.251 (7.606 PME Espanholas e 8.645 PME Portuguesas) e ii) grupo das empresas falidas 285 (145 PME Espanholas e 140 PME Portuguesas).

## **4.2 Descrição das Variáveis**

### **4.2.1 Variável dependente**

Conforme apresentado anteriormente, o objetivo do presente trabalho consiste em desenvolver um modelo de previsão de falência para PME da Península Ibérica e, para tal, a variável dependente será definida sob forma de uma variável *dummy* (“não falidas “ assume o valor de zero e “falidas“ assume o valor de 1) para diferenciar as empresas falidas das empresas não falidas. Portanto, para este estudo foram classificadas como falidas as empresas que de acordo com a base de dados da ORBIS Europe apresentam o *Status Active (insolvency proceedings)*, *Bankruptcy* ou *Dissolved (Bankruptcy)* e como saudáveis ou não falidas as empresas cujo status esta definido como *Active*.

### **4.2.2 Variáveis independentes**

Como variáveis independentes, serão selecionados os rácios económico-financeiros utilizados na literatura existente (p.e, Jackendoff, 1962; Beaver, 1966; Altman, 1968; Ohlson, 1980; Altman e Sabato, 2007; Lacerda e Moro, 2008; Altman et al., 2010). De seguida, esses rácios serão previamente testados com o objetivo de se avaliar sua capacidade para prever eventos como a situação de falência e, conseqüentemente, a sua provável inclusão no modelo como variável independente/explicativa de falência, como já explicado no Capítulo 3 (Metodologia).

**Tabela 2 – Lista das variáveis independentes**

Designação da Variável	Rácio
Dimensão	Log (Ativo Total)
Rácio de Liquidez geral	Ativo Corrente/ Passivo Corrente
Rácio de Liquidez Imediata	Caixa e equivalentes /Passivo Corrente
Rácio do Fundo Maneio	(Ativo Corrente -Passivo Corrente) /Ativo Total
Rácio Inverso da Liquidez Geral	Passivo Corrente/Ativo Corrente
Rácio de Solvabilidade	Capitais próprios /Passivo Total
Rácio de Endividamento Total	Passivo total/Ativo, total
Rotação do Ativo Total	Volume de Negócios/ Ativo Total
Rácio de Endividamento de Curto Prazo	Dividas Financeiras de Curto Prazo /Ativo corrente
Prazo Médio de Inventários em Armazém	(Inventários/Vendas) *365
Rácio de Eficiência do Ativo (CFTA)	Cash Flows/Ativo Total
Rentabilidade do Capitais próprios (ROE)	Resultado Líquido / Capitais próprios
Rentabilidade Operacional do Ativo (ROA)	EBITDA /Ativo Total
Margem do EBITDA	EBITDA /Volume de Negócios
Peso do Imposto sobre Ativo Total	Imposto sobre Rendimento do Exercício / Ativo Total
Rácio de crescimento das vendas	(Vendas ano_t)/(vendas ano_t-1)-1
Rácio de Cobertura de Juros	EBITDA / Juros Pagos
Idade	Log(Ano corrente – Ano da Constituição)
Volatilidade das vendas	Média das Vendas /Desvio Padrão das Vendas
Índice de Concentração Industrial (HHI)	Vendas da Empresa <i>i</i> / Total de Vendas do Setor

**Nota:** A tabela 2 apresenta as variáveis selecionadas para o estudo.

- Dimensão: Logaritmo do Ativo Total – Esta variável é medida pelo logaritmo do total dos ativos, sendo considerada como uma *proxy* da dimensão da empresas (Domenichelli, 2018).

- Rácio de Liquidez Geral: (Ativo Corrente/Passivo Corrente) – Este rácio avalia a capacidade de a empresa cumprir com os seus compromissos financeiros no curto prazo, ou seja, compara o ativo de curto prazo total com o passivo de curto prazo total. O ideal é que o seu valor seja sempre superior a 1, refletindo uma boa situação financeira das empresas no curto prazo. Caso contrário (rácio inferior a 1) prevê-se o aumento da probabilidade de falência da empresa impulsionada pela insuficiência de liquidez para regularizar os passivos exigíveis no curto prazo (Jackendoff, 1962).

- Rácio de Liquidez Imediata: (Caixa e equivalentes /Passivo Corrente) – Obtido pela divisão dos valores financeiros líquidos e o passivo de curto prazo, este rácio expressa a capacidade de a empresa pagar o passivo de curto prazo por meio dos ativos líquidos, i.e, indica se a empresa é capaz de honrar as suas obrigações de curto prazo só com os ativos líquidos disponíveis hoje, sem contar com as contas a receber de

clientes e a venda de algum produto em armazém. Quanto mais elevado for o montante convertido em ativos líquidos, menor será a probabilidade de a empresa falir (Beaver, 1966).

- Inverso da Liquidez Geral:  $(\text{Passivo Corrente}/\text{Ativo Corrente})$  – Sendo o inverso do rácio de liquidez geral, este indicador compara o passivo corrente com o ativo corrente. O que significa que quando o valor das obrigações da empresa até um ano for superior ao ativo corrente, assume-se que a empresa não reflete uma situação financeira saudável no curto prazo. Logo quanto maior o valor deste rácio, maior é a probabilidade de a empresa falir (Ohlson, 1980).

- Rácio do Fundo de Maneio:  $((\text{Ativo Corrente} - \text{Passivo Corrente})/\text{Ativo total})$  – Este rácio exprime a diferença entre o ativo corrente e o passivo corrente em relação ao ativo total, i.e., baseado na regra de equilíbrio financeiro mínimo, expressa a parcela de meios financeiros disponíveis para liquidar as obrigações prestes a vencer de forma regular e contínua, e ainda reflete a margem de segurança para fazer face a qualquer dificuldade que possa surgir. Por norma, esse rácio deve ser igual ou superior a zero e, conseqüentemente, quanto maior o valor do rácio, menor a probabilidade de insolvência (Beaver, 1966; Altman, 1968; Ohlson, 1980).

- Rácio de Solvabilidade:  $(\text{Capitais Próprios}/\text{Passivo Total})$  – Obtido através da relação entre os capitais próprios e os capitais alheios, este rácio procura analisar a estrutura de capitais da empresa. Neste sentido, pressupõe-se a inexistência de problemas financeiros para empresas com um rácio de solvabilidade elevado, enquanto valores baixos representam empresas com um nível de endividamento elevado e com maior probabilidade de falência, ou seja, a quanto mais elevado este rácio, maior a solidez financeira da empresa (Altman, 1968).

- Rácio de Endividamento Total:  $(\text{Passivo Total}/\text{Ativo Total})$  – Relacionado com a estrutura financeira da empresa, este rácio representa a relação entre o passivo total e o ativo total, indicando a proporção do ativo financiada pelo passivo da empresa. Assim sendo, a probabilidade de uma falência iminente é tanto maior quanto mais elevado for este rácio (Ohlson, 1980).

- Rotação do Ativo Total:  $(\text{Volume de Negócios}/\text{Ativo Total})$  – Calculado dividindo o volume de negócios pelo ativo total, este rácio procura avaliar o nível de eficiência da empresa na gestão do seu ativo, isto é, representa o número de unidades monetárias vendidas em relação a cada unidade monetária investida pela empresa.



Desse modo, poder-se-á dizer que quanto mais elevado for esse rácio, maior é a eficiência da empresa e menor o risco de falência (Altman, 1968).

- Endividamento de Curto Prazo: (Dívidas Financeiras de Curto Prazo/Ativo Corrente) – É um rácio que procura avaliar a capacidade da empresa em gerar meios financeiros para honrar as suas obrigações financeiras de curto prazo. Quanto maior o valor desse rácio maior será a incapacidade da empresa em saldar as suas dívidas de curto prazo, sendo desse modo um indiciador de insolvência (Neves & Silva, 1998).

- Rácio Prazo Médio de Inventários em Armazém: ((Inventários/Vendas) \*365) – Considerado como um indicador de funcionamento e eficácia do ciclo operacional, o seu cálculo é fundamental para se conhecer a quantidade e *timing* necessários para se evitar a rotura de stocks, este rácio indica o tempo médio de duração dos inventários em armazém.

- Rácio de Eficiência do Ativo (CFTA): (*Cash Flows*/Ativo Total) – Este rácio reflete a percentagem dos fluxos de caixa em relação aos ativos. Segundo Frank e Urbancic (2005), este rácio avalia a eficiência com que a empresa utiliza os seus ativos na captação de fluxos de caixa provenientes das vendas. Sendo que o seu valor para uma empresa eficiente normalmente é elevado, é considerado um preditor de falência.

- Rentabilidade do Capitais Próprios: (Resultado Líquido/Capitais Próprios) – Designado normalmente por *Return on Equity* ou ROE, calcula-se dividindo os resultados líquidos e os capitais próprios. Este rácio reflete a expectativa de retorno dos investidores quando comparada a investimentos alternativos no mercado e isentos de risco. Segundo Gordon (1962), é um dos indicadores mais importantes da rentabilidade e da eficiência da gestão da empresa, por indicar a taxa de retorno para os acionistas obtida anualmente através dos lucros. Assim, taxas de rentabilidade excessivamente baixas põem em risco a sustentabilidade e continuidade da empresa, provocando assim a escassez de investidores dispostos a arriscar o seu capital.

- Rentabilidade Operacional do Ativo (ROA): (EBITDA/Ativo Total) – Resultante da divisão entre os resultados (antes dos juros, impostos, depreciações e amortizações) e o ativo total, este rácio reflete a capacidade dos ativos das empresas gerarem resultados. Valores elevados para esse rácio representam um bom desempenho operacional da empresa e, por conseguinte, a probabilidade de a empresa entrar em falência é menor (Altman, 1968; Altman, 2017).

- Margem do EBITDA:  $(EBITDA / \text{Volume de Negócios})$  – Também designado por rácio de rentabilidade operacional das vendas, reflete a capacidade do negócio ser rentável, ou seja, expressa o lucro obtido pela empresa por cada unidade vendida, antes de descontar os juros, impostos, depreciações e amortizações. De acordo com este rácio, uma margem EBITDA elevada é sinónimo de uma probabilidade reduzida de a empresa entrar em falência (Altman & Sabato, 2007).

- Peso do Imposto sobre Ativo Total:  $(\text{Imposto sobre Rendimento} / \text{Ativo Total})$  - Este rácio exprime a relação entre os impostos que a empresa paga ao estado e o ativo total, i.e. a diferença entre o rácio do resultado antes de imposto (EBT) sobre o ativo total e o rácio do resultado líquido sobre o ativo total, mencionado na literatura, por exemplo, Beaver (1966).

- Rácio de crescimento das vendas:  $((\text{vendas ano } t / \text{vendas ano } t-1) - 1)$  – Este rácio indica a variação percentual das vendas entre 2 anos consecutivos. . Quanto mais elevado o valor deste rácio, menor a probabilidade de falência.

- Rácio de Cobertura de Juros:  $(EBITDA / \text{Juros pagos})$  – Este rácio avalia a rapidez com que a empresa pode pagar as despesas da dívida. Para uma empresa com resultados mais do que suficientes para cobrir as despesas de financiamento, o seu valor é normalmente superior a 1. Enquanto valores abaixo de 1 significa que a empresa não gera resultados suficientes para cobrir as despesas da dívida e, a longo prazo, devido às dificuldades em financiar-se, a empresa encontrar-se-á numa situação de falência (Cherobim, Junior & Rigo, 2010).

- Idade:  $(\text{Logaritmo} (\text{Ano Corrente} - \text{Ano da Constituição}))$  – Calculado como o logaritmo da diferença entre o ano observado e o ano de constituição da empresa, é indicado para eliminar problemas de heteroscedasticidade relacionados com a escala das variáveis (Domenichelli, 2018).

- Volatilidade das vendas:  $(\text{Média das Vendas} / \text{Desvio Padrão das Vendas})$  – Calcula-se dividindo a média das vendas e o desvio padrão das vendas (para um período de 3 anos). Segundo Lacerda e Moro (2008), a probabilidade de falência diminui ligeiramente com o aumento do rácio crescimento histórico das vendas/volatilidade do crescimento das vendas, confirmando a hipótese de que a significância da tendência nas vendas ou a estabilidade das vendas contribuem para diminuir a probabilidade de falência da empresa.

- Índice de concentração industrial (HHI): (Vendas da Empresa  $i$ /Total de Vendas do Setor) – É um índice calculado por setor que será utilizado para medir o grau de concentração industrial, i.e., quanto menor a concentração, maior o número de empresas por setor e mais concorrencial é o setor. Usando como variável o nível de endividamento, Lacerda e Moro (2008) constataram que a probabilidade de falência diminui quando o índice de concentração de dívida de Herfindahl aumenta e, por isso concluíram que é um forte preditor da situação de falência.

Por fim, é importante referir que todas as variáveis independentes serão submetidas a um processo de *winsorizing* em 1% em cada cauda, de modo a reduzir o impacto dos valores extremos – *outliers* - superiores e inferiores das variáveis, i.e., o método permitirá que os valores extremos das variáveis (percentil 1 e percentil 99) apresentem o mesmo valor no respetivo percentil.

### **4.3 Estatísticas Descritivas**

O tratamento estatístico da amostra bem como os respetivos resultados serão analisados através do Stata Statistical Software, versão 16.1.

As tabelas 3, 4 e 5 apresentam estatísticas descritivas das variáveis independentes selecionadas para o presente estudo, durante o período de 2012 a 2020, após o processo de *winsorizing*.

A tabela 3 apresenta as estatísticas descritivas para a amostra total, a Tabela 4 apresenta estatísticas descritivas para o grupo de países dividida por empresas espanholas e empresas portuguesas, respetivamente, a Tabela 5 apresenta as estatísticas descritivas por Status (PME Saudáveis vs PME Falidas).

**Tabela 3 - Estatísticas Descritivas: Amostra Total**

Variáveis	Nº Obs	Média	D.P.	P.25	P.50	P.75
Dimensão	129.607	6,35	1,28	5,48	6,36	7,25
Liquidez Geral	129.486	2,56	3,87	0,96	1,46	2,52
Liquidez Imediata	125.310	0,65	1,49	0,04	0,16	0,55
Fundo de Maneio	129.607	0,15	0,47	-0,02	0,20	0,43
Inverso da Liquidez Geral	129.577	1,09	1,72	0,40	0,69	1,04
Solvabilidade	129.544	1,24	2,73	0,14	0,44	1,14
Endividamento Total	129.607	0,74	0,52	0,47	0,69	0,88
Rotação do Ativo	129.607	1,53	1,30	0,70	1,19	1,92
Endividamento de Curto Prazo	129.086	0,12	0,28	0,00	0,00	0,13
PM_Inventários em Armazém (dias)	128.981	90,49	221,90	0,02	17,83	78,77
Rácio de Eficiência	129.605	0,05	0,15	0,01	0,05	0,10
Rentabilidade Capitais Próprios	129.592	0,09	0,73	0,01	0,06	0,19
Rentabilidade Operacional do Ativo	129.606	0,07	0,16	0,02	0,06	0,13
Margem do EBITDA	129.606	0,05	0,16	0,02	0,06	0,11
Imposto /Ativo Total	108.544	0,01	0,02	0,00	0,01	0,02
Rácio de crescimento das vendas	128.780	-0,10	0,35	-0,29	-0,10	0,06
Rácio de Cobertura de Juros	112.972	7,04	1,54	5,47	6,82	8,65
Idade	129.251	2,66	0,69	2,30	2,77	3,14
Volatilidade das vendas	128.236	15,60	350,80	3,60	6,22	11,31
Índice de concentração industrial (HHI)	146.209	2,00	0,82	1,00	2,00	3,00

**Nota:** A Tabela 3 apresenta as estatísticas descritivas referentes à amostra total: Número de Observações., Média, Desvio Padrão (D.P.), Percentil 25º (P25), Percentil 50º que corresponde à Mediana (P50), e Percentil 75º (P75). As variáveis são as seguintes: Dimensão, Liquidez Geral, Liquidez Imediata, Fundo de Maneio, Inverso da Liquidez Geral, Solvabilidade, Endividamento Total, Rotação do Ativo, Endividamento de Curto Prazo, PM\_Inventários em Armazém, Rácio de Eficiência, Rentabilidade dos Capitais Próprios, Rentabilidade Operacional do Ativo, Margem do EBITDA, Imposto /Ativo Total, Rácio de crescimento das vendas, Rácio de Cobertura de Juros, Idade, Volatilidade das vendas, Índice de concentração industrial (HHI).

Conforme observado partir de estatísticas apresentadas na Tabela 3. em média. rácio de liquidez geral é 2,56, o que significa que o ativo corrente é superior em relação ao passivo de curto prazo. O Fundo de Maneio representa em média 15% do valor do Ativo Total. A solvabilidade é de 1,24 indicando que, em média, os capitais próprios são superiores ao passivo. A percentagem do Ativo total financiada pelo passivo da empresa corresponde a 74% (Endividamento Total). A rotação do Ativo é em média 1,53, significando que por cada unidade monetária investida no ativo da empresa, o retorno em vendas é de cerca de 1,53 unidades monetárias.

A Tabela 4 apresenta estatísticas descritivas por país – Espanha e Portugal. Podemos observar que relativamente às empresas espanholas, as empresas portuguesas apresentam valores médios superiores para as variáveis Liquidez Geral,

Fundo de Maneio. Endividamento Total, PM\_Inventários em Armazém, Rácio de Cobertura de Juros e Idade, e valores inferiores quanto às variáveis Solvabilidade, Rotação do Ativo, Volatilidade das Vendas e Índice de concentração industrial (HHI).

**Tabela 4 - Estatísticas Descritivas por País (PME Espanholas vs PME Portuguesas)**

Variáveis	PME Espanholas						PME Portuguesas					
	Nº Obs	Média	D.P.	P.25	P.50	P.75	Nº Obs	Média	D.P.	P.25	P.50	P.75
Dimensão	60.156	6,42	1,25	5,56	6,42	7,29	69.451	6,29	1,30	5,42	6,31	7,21
Liquidez Geral	60.097	2,41	3,70	0,89	1,36	2,36	69.389	2,69	4,00	1,01	1,55	2,65
Liquidez Imediata	56.265	0,67	1,47	0,05	0,18	0,60	69.045	0,63	1,51	0,04	0,14	0,51
Fundo de Maneio	60.156	0,11	0,45	-0,05	0,15	0,38	69.451	0,18	0,48	0,01	0,24	0,47
Inverso da Liquidez Geral	60.128	1,19	1,85	0,42	0,74	1,12	69.449	1,01	1,59	0,38	0,65	0,99
Solvabilidade	60.137	1,43	3,05	0,13	0,47	1,30	69.407	1,07	2,42	0,15	0,43	1,03
Endividamento Total	60.156	0,71	0,48	0,43	0,68	0,88	69.451	0,77	0,56	0,49	0,70	0,87
Rotação do Ativo	60.156	1,63	1,38	0,71	1,28	2,10	69.451	1,44	1,22	0,68	1,13	1,76
Endividamento de Curto Prazo	59.850	0,12	0,26	0,00	0,00	0,13	69.236	0,13	0,30	0,00	0,00	0,13
PM_Inventários em Armazém	59.683	88,27	213,38	0,18	18,85	79,88	69.298	92,40	228,96	0,00	17,01	77,80
Rácio de Eficiência	60.155	0,05	0,14	0,01	0,04	0,10	69.450	0,04	0,16	0,01	0,05	0,11
Rentabilidade Capitais Próprios	60.149	0,09	0,72	0,01	0,06	0,19	69.443	0,09	0,74	0,01	0,06	0,19
Rentabilidade Operacional do Ativo	60.156	0,07	0,15	0,02	0,06	0,13	69.450	0,07	0,17	0,02	0,07	0,13
Margem do EBITDA	60.156	0,05	0,16	0,02	0,05	0,10	69.450	0,05	0,17	0,02	0,06	0,12
Imposto /Ativo Total	47.900	0,01	0,03	0,00	0,01	0,02	60.644	0,01	0,02	0,00	0,01	0,02
Rácio de crescimento das vendas	59.796	-0,10	0,34	-0,29	-0,09	0,06	68.984	-0,10	0,36	-0,30	-0,10	0,05
Rácio de Cobertura de Juros	53.455	6,93	1,53	5,47	5,81	8,65	59.517	7,14	1,53	5,47	7,85	8,65
Idade	60.026	2,62	0,62	2,30	2,71	3,04	69.225	2,69	0,75	2,30	2,77	3,18
Volatilidade das vendas	59.607	17,77	485,23	3,60	6,25	11,48	68.629	13,71	159,52	3,59	6,19	11,18
Índice de concentração industrial (HHI)	67.926	2,09	0,83	1,00	2,00	3,00	78.283	1,92	0,80	1,00	2,00	3,00

**Nota:** A Tabela 4 apresenta as estatísticas descritivas por país (Espanha e Portugal): Número de Observações., Média, Desvio Padrão (D.P.), Percentil 25º (P25), Percentil 50º que corresponde à Mediana (P50), e Percentil 75º (P75). As variáveis são as descritas na Tabela 3 e Subcapítulo 4.2.2.

A Tabela 5 apresenta estatísticas descritivas em função do *status* – falidas versus não falidas. Observa-se que os valores médios das variáveis Liquidez Geral e Imediata, Fundo de Maneio, Solvabilidade, Rotação do Ativo, Margem do EBITDA, Rentabilidade dos Capitais Próprios, Rentabilidade Operacional do Ativo, Rácio de Cobertura de Juros são superiores nas empresas saudáveis em relação às empresas falidas, e inferiores quanto às variáveis Inverso da Liquidez Geral, Endividamento Total e de curto prazo, PM\_Inventários em Armazém, Idade, Volatilidade das vendas e Índice de concentração industrial (HHI).

**Tabela 5 - Estatística Descritiva por Status (PME Saudáveis vs PME Falidas)**

Variáveis	PME Saudáveis						PME Falidas					
	Nº Obs	Média	D.P.	P.25	P.50	P.75	Nº Obs	Média	D.P.	P.25	P.50	P.75
Dimensão	127.540	6,34	1,27	5,48	6,35	7,24	2.067	6,90	1,36	5,94	6,96	7,92
Liquidez Geral	127.420	2,57	3,89	0,96	1,46	2,53	2.066	1,90	2,57	0,90	1,26	1,96
Liquidez Imediata	123.321	0,65	1,50	0,04	0,16	0,56	1.989	0,30	0,91	0,02	0,07	0,23
Fundo de Maneio	127.540	0,15	0,47	-0,02	0,20	0,43	2.067	0,09	0,44	-0,05	0,13	0,34
Inverso da Liquidez Geral	127.510	1,09	1,72	0,39	0,68	1,04	2.067	1,21	1,80	0,51	0,79	1,11
Solvabilidade	127.477	1,25	2,75	0,14	0,45	1,15	2.067	0,58	1,43	0,06	0,24	0,59
Endividamento Total	127.540	0,74	0,52	0,46	0,69	0,88	2.067	0,85	0,49	0,63	0,81	0,94
Rotação do Ativo	127.540	1,53	1,30	0,70	1,19	1,92	2.067	1,33	1,13	0,63	1,04	1,69
Endividamento de Curto Prazo	127.025	0,12	0,28	0,00	0,00	0,13	2.061	0,18	0,32	0,00	0,05	0,25
PM_Inventários em Armazém	126.930	90,20	221,63	0,00	17,62	78,18	2.051	108,37	237,13	3,35	32,58	116,76
Rácio de Eficiência	127.538	0,05	0,15	0,01	0,05	0,11	2.067	0,00	0,16	-0,01	0,02	0,06
Rentabilidade Capitais Próprios	127.525	0,09	0,73	0,01	0,06	0,19	2.067	0,04	0,91	-0,04	0,03	0,17
Rentabilidade Operacional do Ativo	127.539	0,07	0,16	0,02	0,07	0,13	2.067	0,02	0,17	0,00	0,04	0,08
Margem do EBITDA	127.539	0,05	0,16	0,02	0,06	0,11	2.067	0,01	0,19	-0,01	0,04	0,08
Imposto /Ativo Total	106.910	0,01	0,02	0,00	0,01	0,02	1.634	0,01	0,02	0,00	0,00	0,01
Rácio de crescimento das vendas	126.750	-0,10	0,35	-0,29	-0,09	0,06	2.030	-0,20	0,37	-0,42	-0,17	-0,02
Rácio de Cobertura de Juros	111.049	7,06	1,54	5,47	6,95	8,65	1.923	6,25	1,32	5,47	5,47	6,67
Idade	127.194	2,66	0,69	2,30	2,77	3,14	2.057	2,68	0,76	2,20	2,77	3,22
Volatilidade das vendas	126.205	15,52	349,54	3,60	6,23	11,31	2.031	20,32	421,80	3,07	5,68	11,47
Índice de concentração industrial (HHI)	143.856	1,99	0,81	1,00	2,00	3,00	2.353	2,71	0,64	3,00	3,00	3,00

**Nota:** A Tabela 5 apresenta as estatísticas descritivas por Status (falidas e não falidas): Número de Observações., Média, Desvio Padrão (D.P.), Percentil 25º (P25), Percentil 50º que corresponde à Mediana (P50), e Percentil 75º (P75). As variáveis são as descritas na Tabela 3 e Subcapítulo 4.2.2.

As Tabelas 6, 7 e 8 apresentam o resumo das diferentes segmentações da amostra por anos, por código da atividade económica (CAE), por país e em relação ao *status* da empresa (saudáveis vs falidas).

Na Tabela 6, referente à descrição da amostra por anos, podemos observar uma maior representatividade de empresas nos anos 2016 e 2017.

**Tabela 6 - Descrição da Amostra por Ano**

Ano	Status-Saudáveis # Observações	Status-Falidas # Observações	Amostra Total
2012	16.230	283	16.513
2013	16.230	285	16.515
2014	16.224	285	16.509
2015	16.236	285	16.521
<b>2016</b>	<b>16.237</b>	<b>285</b>	<b>16.522</b>
<b>2017</b>	<b>16.242</b>	<b>284</b>	<b>16.526</b>
2018	16.235	284	16.519
2019	16.221	283	16.504
2020	14.001	79	14.080
<b>Total</b>	<b>143.856</b>	<b>2.353</b>	<b>146.209</b>

**Nota:** A Tabela 6 apresenta o número de observações da amostra por ano.

A Tabela 7 apresenta da descrição da amostra por país, onde podemos observar que Portugal é o país com maior representatividade de empresa saudáveis. enquanto a Espanha concentra o maior número de empresas falidas.

**Tabela 7 - Descrição da Amostra por País**

País	Status-Saudáveis # Observações		Status-Falidas # Observações		Amostra Total	
	#Observações	#Empresas	#Observações	#Empresas	#Observações	#Empresas
Espanha	66.734	7.606	1.192	145	67.926	7.751
Portugal	77.122	8.645	1.161	140	78.283	8.785
<b>Total</b>	143.856	16.251	2.353	285	146.209	16.536

**Nota:** A Tabela 7 apresenta o número de observações e de empresas por país, e *Status*:saudáveis e falidas e amostra total.

Na Tabela 8 observamos a descrição da amostra por Código de Atividade Económica (CAE) subdividida em 14 categorias industriais. que foram obtidas mediante a redução do CAE<sup>7</sup> (ou NACE) a 2 dígitos. De acordo com os resultados obtidos, a indústria 5 - Comércio por grosso. retalho e serviços de reparação - concentra 26% da amostra total, seguida da indústria 2 - Indústrias extrativas e transformadoras - e indústria 4 - Construção - com uma representatividade de 18% e 16%, respetivamente. Por outro lado, a indústria 2 representa 36% das empresas falidas, seguida da indústria 5 - Comércio por grosso. retalho e serviços de reparação

<sup>7</sup> O código de atividade recolhido na base de dados ORBIS Europe é o NACE, que está altamente correlacionado com o CAE.

- com 24% das empresas falidas.

**Tabela 8 - Descrição da Amostra por CAE**

Indústria	Amostra Total		%	Status-Falida		%
	#Observações	# Empresas		(#Emp)	#Observações	
1	4.587	518	3%	25	3	1%
2	25.691	2.904	18%	852	103	36%
3	826	93	1%	9	1	0%
4	23.887	2.703	16%	216	26	9%
5	37.957	4.288	26%	565	68	24%
6	13.282	1.500	9%	203	25	9%
7	13.467	1.531	9%	197	24	8%
8	3.015	342	2%	8	1	0%
9	7.528	851	5%	82	10	4%
10	6.161	697	4%	73	9	3%
11	1.942	220	1%	25	3	1%
12	4.443	501	3%	41	5	2%
13	1.756	200	1%	24	3	1%
14	1.667	188	1%	33	4	1%
<b>Total</b>	<b>146.209</b>	<b>16.536</b>	<b>100%</b>	<b>2.353</b>	<b>285</b>	<b>100%</b>

**Nota:** A Tabela 8 apresenta a descrição da amostra por indústria, considerando a amostra total (número de observações e número de empresas) e *status* - falidas (número de observações e número de empresas falidas).

Por último, a Tabela 9 apresenta os resultados da amostra por região (NUTS II)<sup>8</sup>, onde podemos observar que para Portugal, as regiões com maior representatividade de empresas falidas são Região do Norte, Região do Centro e Área Metropolitana de Lisboa, com uma concentração de 53%, 19.4% e 18.9%, respetivamente. Já em relação às empresas espanholas, as regiões com maior representatividade de empresas falidas são a Catalunha com 26%, seguida das regiões Comunidade Valenciana e Andaluzia, ambas com 14%, e a Comunidade Autónoma de Madrid com uma concentração de 13%.

<sup>8</sup> NUTS II: corresponde a Nomenclatura das Unidades Territoriais para Fins Estatísticos.



**Tabela 9 - Descrição da Amostra por NUTS II**

<b>Portugal</b>				
<b>Região NUTS II</b>	<b>Amostra Total</b>		<b>PME Falidas</b>	
	<b>#Observações</b>	<b># Empresas</b>	<b>#Observações</b>	<b># Empresas</b>
Região do Centro	18.736	2.072	225	27
Região do Norte	32.052	3.520	615	74
Alentejo	4.465	498	23	3
Algarve	3.199	354	43	5
Área Metropolitana de Lisboa	17.553	1.950	219	27
Região Autónoma dos Açores	921	101	18	2
Região Autónoma da Madeira	1.357	150	18	2
<b>Total</b>	<b>78.283</b>	<b>8.645</b>	<b>1.161</b>	<b>140</b>
<b>Espanha</b>				
<b>Região NUTS II</b>	<b>Amostra Total</b>		<b>PME Falidas</b>	
	<b>#Observações</b>	<b># Empresas</b>	<b>#Observações</b>	<b># Empresas</b>
Galiza	3.611	408	50	6
Astúrias	420	48	24	3
Cantábria	475	54	0	0
País Basco	3.439	389	42	5
Navarra	672	76	0	0
La Rioja	279	32	9	1
Aragão	1.127	128	33	4
Comunidade Autónoma de Madrid	6.182	703	151	18
Castela e Leão	2.981	341	39	5
Castela-Mancha	2.102	240	40	5
Estremadura	1.546	175	17	2
Catalunha	14.863	1.701	310	38
Comunidade Valenciana	12.802	1.463	165	20
Ilhas Baleares	997	116	24	3
Andaluzia	12.418	1.421	164	20
Múrcia	1.864	211	17	2
Ceuta	18	2	0	0
Melilla	27	3	0	0
Ilhas Canárias	2.103	240	107	13
<b>Total</b>	<b>67.926</b>	<b>7.751</b>	<b>1.192</b>	<b>145</b>

**Nota:** A Tabela 9 apresenta a descrição da amostra por NUTS II: considerando a amostra total (número de observações e número de empresas) e *status* - falidas (número de observações e número de empresas falidas). As Unidades Territoriais para Fins Estatísticos - NUTS - consideradas na referida Tabela são representativas dos países que formam a Península Ibérica (Portugal e Espanha).

#### 4.4 Análise da Correlação

Considerada uma análise preliminar em relação à análise empírica, a análise da correlação de Pearson consiste na estimação do grau de correlação entre as variáveis, i.e., estima-se o grau de correlação com o objetivo de se observar como as variáveis se relacionam entre si. Segundo Marôco (2003), a correlação destina-se exclusivamente a analisar o grau de relação entre duas variáveis, sem qualquer implicação de causa efeito entre elas e é representada exclusivamente por um coeficiente de correlação ( $\rho$ ). O referido coeficiente mede o grau de correlação entre duas variáveis quantitativas e apresenta valores situados entre -1 e 1, inclusive. O sinal do coeficiente indica-nos a direção dessa correlação, i.e., se as variáveis estão positivas ou negativamente correlacionadas, embora a sua interpretação<sup>9</sup> não seja consensual.

Na Tabela 10 podemos observar os resultados da estimação dos coeficientes da correlação de Pearson e os níveis de significância entre todas as variáveis selecionadas e descritas no Subcapítulo 4.2.2. Com base nos resultados é possível verificar que a maioria das variáveis revela estar correlacionada de forma significativa com a variável dependente, à exceção da Idade e da Volatilidade das Vendas.

De acordo com o esperado, a nossa variável dependente, que assume 1 quando a empresa se encontra em situação de falência, está negativamente correlacionada com as variáveis Liquidez Geral e Imediata, Fundo de Maneio, Solvabilidade, Rotação do Ativo, Rácio de Eficiência, Rentabilidade dos Capitais Próprios, Rentabilidade Operacional do Ativo, Margem do EBITDA, Imposto /Ativo Total, Crescimento das Vendas e Rácio de Cobertura. Por outro lado, está positivamente correlacionada com o Inverso da Liquidez, Endividamento Total e de Curto Prazo, PM\_Inventários em Armazém e Índice de Concentração Industrial (HHI).

---

<sup>9</sup> Segundo Hinkle, Wiersma & Jurs (2003), a intensidade da correlação pode ser interpretada da seguinte forma: ]0,90; 1] / ]-0,90; -1]: Correlação muito alta positiva/negativa; ]0,70; 0,90[ / ]-0,70; -0,90[: Correlação alta positiva/negativa; ]0,50; 0,70[ / ]-0,50; -0,70[: Correlação moderada positiva/negativa; 0,30; 0,50[ / ]-0,30; -0,50[: Correlação baixa positiva/negativa; [0; 0,30[ / [0; -0,30[: Correlação insignificante. Segundo Franzblau (1958), a intensidade da Correlação de Pearson pode ser interpretada da seguinte forma: sem correlação (menor que 0,2); fraca correlação (entre 0,2 e 0,4); correlação moderada (entre 0,4 e 0,6); forte correlação (entre 0,6 e 0,8) e muito forte correlação (maior que 0,8).

**Tabela 10 - Matriz de Correlação de Pearson**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
<b>Falida (1)</b>	1																				
<b>Dimensão (2)</b>	0,0546*	1																			
<b>LG (3)</b>	-0,0217*	0,0316*	1																		
<b>LI (4)</b>	-0,0295*	-0,0421*	0,7239*	1																	
<b>FM (5)</b>	-0,0164*	0,1732*	0,4534*	0,3300*	1																
<b>Inverso LG (6)</b>	0,0089*	-0,1358*	-0,2712*	-0,1927*	-0,7335*	1															
<b>Solvabilidade (7)</b>	-0,0305*	0,0999*	0,6351*	0,6177*	0,3439*	-0,2060*	1														
<b>Endividamento (8)</b>	0,0275*	-0,3052*	-0,3020*	-0,2840*	-0,7371*	0,5204*	-0,4614*	1													
<b>RA (9)</b>	-0,0197*	-0,4335*	-0,1298*	-0,0327*	-0,1783*	0,0909*	-0,1292*	0,2461*	1												
<b>Endividamento CP (10)</b>	0,0263*	0,1384*	-0,1854*	-0,1503*	-0,3088*	0,3699*	-0,1253*	0,1130*	-0,0584*	1											
<b>PMI (11)</b>	0,0102*	0,1631*	0,1188*	-0,0641*	0,0753*	-0,0713*	-0,0126*	0,0305*	-0,2686*	-0,0436*	1										
<b>R. Eficiência (12)</b>	-0,0415*	0,1251*	0,0621*	0,1170*	0,3240*	-0,2492*	0,1114*	-0,4310*	0,0374*	-0,0457*	-0,1338*	1									
<b>RCP (13)</b>	-0,0086*	-0,0313*	-0,0134*	0,0066*	-0,0220*	0,0068*	-0,0181*	0,0387*	0,0664*	-0,0172*	-0,0443*	0,0925*	1								
<b>ROA (14)</b>	-0,0385*	0,1062*	0,0509*	0,1100*	0,3120*	-0,2408*	0,0931*	-0,4018*	0,0687*	-0,0337*	-0,1474*	0,9820*	0,1012*	1							
<b>Margem EBITDA (15)</b>	-0,0349*	0,1863*	0,0670*	0,1139*	0,2498*	-0,1903*	0,1158*	-0,3488*	-0,0714*	0,0007	-0,1781*	0,6966*	0,0746*	0,7064*	1						
<b>Imposto/Ativo (16)</b>	-0,0381*	-0,1018*	0,0546*	0,1443*	0,1767*	-0,1255*	0,0929*	-0,1515*	0,2580*	-0,1036*	-0,1702*	0,5861*	0,1747*	0,6750*	0,4237*	1					
<b>%Δ vendas (17)</b>	-0,0359*	0,0069*	0,0170*	-0,0142*	0,0708*	-0,0643*	-0,0021	-0,0397*	-0,0660*	-0,0427*	0,0499*	0,0051	0,0012	0,0025	-0,0022	-0,0156*	1				
<b>Cobertura (18)</b>	-0,0679*	0,0120*	0,0995*	0,1817*	0,1904*	-0,1372*	0,2027*	-0,2730*	0,1181*	-0,1504*	-0,1722*	0,5099*	0,1174*	0,4896*	0,4014*	0,4336*	0,0034	1			
<b>Idade (19)</b>	0,0039	0,2891*	0,0890*	0,0489*	0,0952*	-0,0550*	0,1283*	-0,1242*	-0,2182*	0,0348*	0,1143*	-0,0535*	-0,0718*	-0,0676*	-0,005	-0,1044*	-0,0292*	-0,0381*	1		
<b>SD vendas (20)</b>	0,0017	0,0068*	-0,002	0	-0,0004	-0,0008	0,0002	-0,0025	0,0033	0,0025	-0,005	0,0037	-0,0012	0,0009	0,0035	0,0018	0,0059	0,0027	0,0117*	1	
<b>HHI (21)</b>	0,1111*	0,0123*	-0,0063*	-0,0385*	-0,0305*	0,0187*	-0,0362*	0,0571*	-0,0351*	-0,0139*	0,1029*	-0,0248*	0,0165*	-0,0204*	-0,0633*	0,0221*	0,1026*	-0,0312*	-0,1761*	-0,0063*	1

**Nota:** A Tabela 10 apresenta os coeficientes de correlação de Pearson. \* Significância estatística ao nível de 5% ou inferior. As variáveis são as seguintes: (1) Status – falidas vs não falidas, (2) Dimensão, (3) Liquidez Geral, (4) Liquidez Imediata, (5) Fundo de Maneio, (6) Inverso da Liquidez Geral (7) Solvabilidade, (8)Endividamento Total, (9) Rotação do Ativo, (10)Endividamento de Curto Prazo, (11) PM\_Inventários em Armazém, (12)Rácio de Eficiência, (13) Rentabilidade dos Capitais Próprios, (14)Rentabilidade Operacional do Ativo, (15)Margem do EBITDA, (16)Imposto /Ativo Total, (17)Rácio de crescimento das vendas, (18)Rácio de Cobertura de Juros, (19) Idade, (20)Volatilidade das vendas, (21) Índice de concentração industrial (HHI).

## 4.5 Análise Univariada

A Análise Univariada foi a metodologia estatística utilizada nos primeiros estudos empíricos sobre modelos de previsão de falência. Dos estudos realizados tendo como base a respetiva metodologia destacam-se os estudos de autores como FitzPatrick (1932), Smith e Winakor (1935) e Merwin (1942), que testaram individualmente vários rácios económico-financeiros com o objetivo de identificar e analisar os que melhor discriminam o grupo de empresas falidas face ao grupo de empresas saudáveis. Mais tarde Beaver (1966), utilizando a mesma metodologia, realizou o seu estudo que é considerado um dos mais importantes no conjunto de estudos que adotou a análise univariada. Embora Beaver (1966) tenha obtido resultados satisfatórios no seu estudo, o autor levantou a hipótese de que a capacidade preditiva do modelo pode melhorar caso os rácios sejam testados em simultâneo e não individualmente, impulsionando desta forma o trabalho Altman (1968), que utilizou como metodologia a Análise Discriminante Múltipla (ADM).

Assim, como análise preliminar, para além da análise de correlação, conduziu-se uma análise univariada para se testar as diferenças às médias e medianas entre o grupo de empresas falidas e o grupo das empresas saudáveis e entre países (Espanha vs Portugal). Desta forma, efetuaremos um teste paramétrico às médias. o *t* test. e um teste não paramétrico às medianas. designado por Wilcoxon-Mann-Whitney rank sum test.

A Tabela 11 apresenta os resultados dos referidos testes. i.e., a comparação das médias e das medianas das variáveis quantitativas dos dois grupos independentes. Primeiro em função do país - Espanha vs Portugal - e depois em função do *Status* - falência vs não falidas.

Os resultados obtidos revelam que em relação à análise em função do país (Espanha vs Portugal), a maioria das variáveis apresenta diferenças entre as médias e medianas estatisticamente significativas ao nível de 1%, à exceção da variável Volatilidade das vendas que é estatisticamente significativa ao nível de 5% (no caso das diferenças entre médias) e das variáveis Rentabilidade dos Capitais Próprios (no caso das diferenças entre médias e medianas) e na Rentabilidade Operacional do Ativo e Rácio de crescimento das vendas (no caso das diferenças entre médias) que não apresentam significância estatística.

Relativamente ao *Status* (falidas vs não falidas), podemos observar que, à exceção das variáveis Idade e Volatilidade das vendas que não apresentam significância estatística (no caso das diferenças entre médias) e da variável Rentabilidade Capitais Próprios que a diferenças entre médias é estatisticamente significativa ao nível de 5%, a diferença entre as médias e as medianas das restantes variáveis apresentam significância estatística ao nível de 1%.

**Tabela 11 - Teste às Médias e as Medianas por País e Status**

Variável	País		PME Falidas	
	Diferenças entre médias <i>P-value</i>	Diferenças entre medianas <i>P-value</i>	Diferenças entre médias <i>P-value</i>	Diferenças entre medianas <i>P-value</i>
Dimensão	0,0000	0,0000	0,0000	0,0000
Liquidez Geral	0,0000	0,0000	0,0000	0,0000
Liquidez Imediata	0,0000	0,0000	0,0000	0,0000
Fundo de Maneio	0,0000	0,0000	0,0000	0,0000
Inverso da Liquidez Geral	0,0000	0,0000	0,0023	0,0000
Solvabilidade	0,0000	0,0000	0,0000	0,0000
Endividamento Total	0,0000	0,0000	0,0000	0,0000
Rotação do Ativo	0,0000	0,0000	0,0000	0,0000
Endividamento de Curto Prazo	0,0000	0,0000	0,0000	0,0000
PM_Inventários em Armazém	0,0008	0,0003	0,0006	0,0000
Rácio de Eficiência	0,0000	0,0000	0,0000	0,0000
Rentabilidade Capitais Próprios	0,7035	0,2451	0,0120	0,0000
Rentabilidade Operacional do Ativo	0,3416	0,0000	0,0000	0,0000
Margem do EBITDA	0,0000	0,0000	0,0000	0,0000
Imposto /Ativo Total	0,0000	0,0000	0,0000	0,0000
Rácio de crescimento das vendas	0,8037	0,0000	0,0000	0,0000
Rácio de Cobertura de Juros	0,0000	0,0000	0,0000	0,0000
Idade	0,0000	0,0000	0,2017	0,0096
Volatilidade das vendas	0,0505	0,0054	0,6098	0,0000
Índice de concentração industrial (HHI)	0,0000	0,0000	0,0000	0,0000

**Nota:** A Tabela 11 apresenta o *p-value* dos testes às médias e às medianas por grupos: País (Espanha vs Portugal) e Status (PME falidas vs PME Saudáveis). As variáveis são as descritas no Subcapítulo 4.2.2.

Comparando informação das Tabela 4 – Estatísticas Descritivas por grupos (PME Espanholas vs PME Portuguesas) com a informação apresentada na Tabela 11, observa-se que o rácio de Solvabilidade e Rotação do Ativo são superiores nas empresas espanholas, o que nos leva a concluir as empresas portuguesas enfrentam maiores dificuldades em solver as suas obrigações de médio e longo prazo e são menos eficientes do que as empresas espanholas. Em suma, estes resultados sugerem que as empresas portuguesas são mais vulneráveis e mais propensas a eventos de falência do que as espanholas.

## 5. Análise Multivariada

A metodologia Multivariada em modelos de previsão de falência foi utilizada pela primeira em 1968 por Altman, com o objetivo de ultrapassar as restrições impostas pela metodologia univariada no desenvolvimento de modelos de previsão de falência, desenvolvendo desse modo o seu famoso modelo designado por Z-score, inspirado pelo trabalho de Beaver (1966). Os métodos multivariados diferem dos univariados no sentido de analisarem as diferenças entre grupos bem definidos à partida com base num conjunto relevante de características dos elementos que compõe esses grupos.

Para este estudo, a metodologia Multivariada será conduzida mediante a utilização dos diferentes métodos estatísticos, detalhados no Capítulo 3, para o processo de seleção das variáveis a incluir no modelo final de previsão de falência, como é o caso da análise discriminante múltipla (ADM), a análise logística e o método *Propensity Score Matching* (PSM) para garantir a segurança na seleção das variáveis e ainda reduzir o enviesamento causado pela desproporcionalidade da amostra.

### 5.1 Análise Discriminante Multivariada

Como o objetivo principal deste trabalho é desenvolver um modelo de previsão de falência com recurso a análise logística, a Análise Discriminante Múltipla (ADM), será utilizada como metodologia complementar para auxiliar na seleção das variáveis que irão ser incluídas nesse modelo. Deste modo, o processo de implementação da ADM apresenta-se estruturado da seguinte forma: i) classificação das empresas em falidas e não falidas ii) seleção das variáveis. iii) análise de matriz de correlação de Pearson e iv) análise da variância (ANOVA).

A primeira fase da ADM consiste na definição da variável dependente, uma variável *dummy* (designada por *Status*) que classifica as empresas falidas com o valor de 1 e as não falidas com o valor de zero de acordo com o conjunto de variáveis independentes apresentadas anteriormente (ver Subcapítulo 4.2.2).

A Tabela 12, denominada por *Confusion Table*, é o resultado da implementação da primeira etapa a ADM, com base na classificação das empresas falidas/saudáveis (identificadas pela variável dependente) e na classificação produzida pelas variáveis

independentes usadas neste estudo. E tendo em conta o resultado da referida classificação, observamos que 75,05% (72,75%) das empresas classificadas como falidas (não falidas) estão corretamente classificadas, enquanto que 24,95% (27,25%) das empresas classificadas como falidas (não falidas) deveriam estar classificadas como não falidas (falidas) de acordo com as variáveis independentes seleccionadas para este estudo.

**Tabela 12 – Classificação das empresas conforme o seu status: *Confusion Table***

<i>True Status</i>	Classificação		Total
	0	1	
0	65.706 72,75%	24.617 27,25%	90.23 100%
1	0.358 24,95%	1.077 75,05%	1.435 100%
Total	66.064 72,00%	25.694 28%	91.758 100%

**Nota:** A Tabela 12 apresenta a *Confusion Table*, que resulta da classificação das empresas consideradas falidas com valor (1) e as empresas consideradas como saudáveis com o valor zero (0), de acordo com as variáveis independentes seleccionadas para este estudo apresentadas detalhe no Subcapítulo 4.2.2.

Resumindo, a *Confusion Table* permite avaliar a capacidade discriminante do modelo com base em todas as variáveis usadas na análise.

Conforme a classificação proposta por Ohlson (1980) e apresentada na Tabela 13, são classificados como Erros do Tipo I os falsos positivos que resultam da classificação de empresas saudáveis como falidas, e Erros do Tipo II os falsos negativos que resultam da classificação de empresas falidas como saudáveis.

**Tabela 13 - Erros tipos I e II decorrentes da classificação das empresas**

Previsão do Resultado	Verdadeiro Positivo Empresas Falidas(y=1)	Falso Positivo Erro tipo I: empresas saudáveis classificadas como falidas
	Falso Negativo Erro tipo II: empresas falidas classificadas como saudáveis	Verdadeiro Negativo Empresas saudáveis (y=0)

**Nota:** A Tabela 13 apresenta a classificação dos Erros Tipo I e II de acordo com a interpretação feita por Ohlson (1980).

Reduzir a probabilidade de erro nos modelos de previsão de falência é importante porque melhora a precisão do modelo e, além disso, uma má classificação

tem custos elevados, principalmente quando empresas falidas são classificadas como não falidas. ou seja, Erro do Tipo II. Deste modo. na fase seguinte da ADM serão selecionadas as variáveis que permitem diminuir a percentagem de erro e que revelam uma elevada capacidade preditiva para identificar potenciais empresas falidas. Para tal, serão conduzidas as seguintes análises que constituem a ADM, nomeadamente, Análise dos Coeficientes *Unstandardized* e *Standardized* incluídos na função discriminante, Análise da Correlação de Pearson e Análise ANOVA. que funcionam como análise complementar à análise logística na seleção de variáveis para o modelo final de previsão e falência.

Os coeficientes *unstandardized* representam o peso de cada variável na classificação realizada na primeira fase da ADM de empresas falidas/não falidas, independentemente do seu grupo. sendo por isso considerados para o cálculo do *score* na função discriminante (ver equação 1 no Subcapítulo 3.1). Esta análise dos coeficientes *unstandardized* deve ser observada como uma ferramenta para definir a função discriminante. e não como o método de seleção de variáveis *per se*. Com a definição da função discriminante, espera-se que os seus preditores maximizem a distância entre os grupos (falidas/não, falidas).

Apesar da análise dos coeficientes *unstandardized* ser considerada uma ferramenta para a definição da função discriminante, esta apresenta limitações em relação, por exemplo, à escala das variáveis. É disso exemplo as variáveis Dimensão e Idade que são calculadas como logaritmo, enquanto as restantes variáveis independentes são rácios calculados numa escala diferente. Dado que esta análise aos coeficientes da função discriminante é baseada no peso de cada variável na discriminação entre grupos, logo esta análise pode conduzir a um enviesamento na seleção das variáveis mais relevantes, i.e., coeficientes mais “pesados” têm maior peso na explicação da situação de falência. De forma a mitigar esse problema, recorre-se ao processo de estandardização dos coeficientes que corresponde à normalização da escala das variáveis, o que contribui para a clarificação da análise da capacidade discriminante de cada variável.

Os coeficientes *standardized* devem ser interpretados em função do seu valor absoluto (i.e., ignorando o sinal dos coeficientes).



**Tabela 14 - Coeficientes da Função Discriminante**

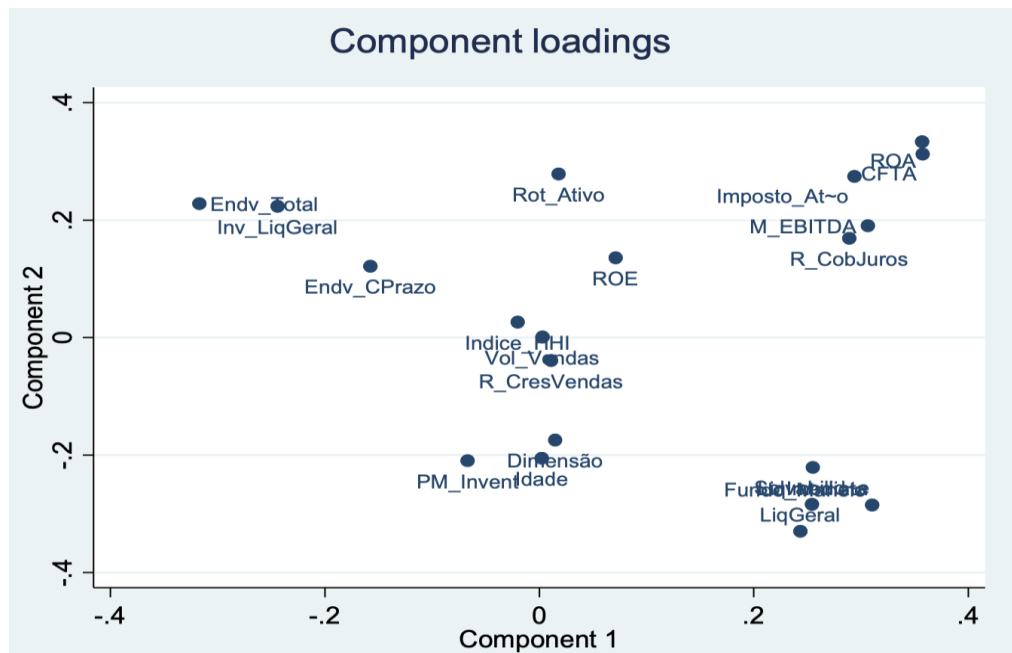
Variáveis	Coeficientes <i>Unstandardized</i>	Coeficientes <i>Standardized</i>
Dimensão	0,3326	<b>0,3941</b>
Liquidez Geral	-0,0140	-0,0459
Liquidez Imediata	-0,0174	-0,0218
Fundo de Maneio	0,5874	<b>0,2214</b>
Inverso da Liquidez Geral	-0,0178	-0,0243
Solvabilidade	-0,0161	-0,0362
Endividamento Total	0,5235	<b>0,2018</b>
Rotação do Ativo	0,0351	0,0402
Endividamento de Curto Prazo	0,3088	0,0855
PM_Inventários em Armazém	-0,0007	-0,1432
Rácio de Eficiência	-3,6981	<b>-0,4211</b>
Rentabilidade Capitais Próprios	-0,0279	-0,0177
Rentabilidade Operacional do Ativo	3,4835	<b>0,4469</b>
Margem do EBITDA	-0,5510	-0,0726
Imposto /Ativo Total	-5,7980	-0,1219
Rácio de crescimento das vendas	-0,8974	<b>-0,3086</b>
Rácio de Cobertura de Juros	-0,2032	<b>-0,3095</b>
Idade	0,0926	0,0626
Volatilidade das vendas	0,0000	-0,0025
Índice de concentração industrial (HHI)	0,9589	<b>0,7720</b>
Constante	-3,3202	-

**Nota:** A Tabela 14 apresenta os coeficientes *unstandardized* e os coeficientes *standardized* que correspondem à 2ª etapa da ADM. As variáveis são as descritas no Subcapítulo 4.2.2.

Deste modo, as variáveis deverão ser ordenadas em função do seu peso (mais uma vez, ignorando o seu sinal). De acordo com os resultados apresentados na Tabela 14 relativamente aos dos coeficientes *standardized*, podemos concluir que as variáveis Dimensão, Fundo de Maneio, Endividamento Total, Rácio de Eficiência, Rentabilidade Operacional do Ativo, Rácio de crescimento das vendas, Rácio de Cobertura de Juros e Índice de Concentração Industrial (HHI) são as que demonstraram maior peso na discriminação entre empresas falidas/não falidas. Importa salientar que, por norma, o *cutoff point* considerado razoável como determinante na identificação de variáveis com maior ou menor poder discriminatório é de 30% (Sage. 2015. p. 600). No entanto, assumindo que o *cutoff point* não é estático e que não existe consenso absoluto na definição do mesmo, 20% foi o *cutoff point* definido para este estudo.

Para sustentar a decisão de seleção das variáveis, nesta fase é fundamental observar a representação gráfica dos coeficientes para se verificar se existe um elevado grau de colinearidade entre variáveis (o que não deve acontecer), sendo que a sua presença revelaria um elevado grau de correlação entre as variáveis (i.e., superior ao desejável).

**Gráfico 1 – Representação Gráfica dos Coeficientes *Standardized***



**Nota:** O Gráfico 1 apresenta os coeficientes *standardized*. As variáveis são as descritas no Subcapítulo 4.2.2.

Efetuada uma comparação entre o Gráfico 1 com os coeficientes *standardized* da Tabela 14, verifica-se que variáveis Rentabilidade Operacional do Ativo e Rácio de Eficiência, Fundo de Maneio, Solvabilidade, Liquidez Geral e Liquidez Imediata encontram-se sobrepostas, o que o que dificulta a diferenciação entre grupos, alertando também para um elevado grau de correlação entre as variáveis. Esta análise será levada em consideração na seleção das variáveis para o modelo final.

Com o objetivo de reforçar os resultados obtidos na análise dos coeficientes *standardized*, a etapa seguinte da ADM é a Análise da Correlação de Pearson entre grupos em função das variáveis explicativas, designada por *structure matrix*. Os coeficientes, tal como na análise anterior, são interpretados em função do seu valor absoluto e representam o grau de correlação (de Pearson) entre o peso de cada variável e a função discriminante. A Tabela 15 apresenta o resultado da análise da

correlação de Pearson.

**Tabela 15 - Coeficientes de Pearson — Structure Matrix**

Variáveis	Coeficientes
Dimensão	<b>0,3644</b>
Liquidez Geral	-0,1248
Liquidez Imediata	-0,1836
Fundo de Maneio	-0,0778
Inverso da Liquidez Geral	0,0455
Solvabilidade	-0,1863
Endividamento Total	0,1878
Rotação do Ativo	-0,1204
Endividamento de Curto Prazo	0,1747
PM_Inventários em Armazém	0,0730
Rácio de Eficiência	<b>-0,2735</b>
Rentabilidade Capitais Próprios	-0,0691
Rentabilidade Operacional do Ativo	<b>-0,2490</b>
Margem do EBITDA	<b>-0,2101</b>
Imposto /Ativo Total	<b>-0,2241</b>
Rácio de crescimento das vendas	<b>-0,2311</b>
Rácio de Cobertura de Juros	<b>-0,4366</b>
Idade	0,0325
Volatilidade das vendas	-0,0153
Índice de concentração industrial (HHI)	<b>0,7312</b>

**Nota:** A Tabela 15 apresenta a análise dos coeficientes de Pearson que avaliam a correlação entre cada variável explicativa e a função discriminante. As variáveis são as descritas no Subcapítulo 4.2.2.

Mais uma vez, considerando um *cutoff point* de 20% (conforme explicado na análise da Tabela 14), os coeficientes que apresentam maior grau de correlação são: Dimensão, Rácio de Eficiência, Rentabilidade Operacional do Ativo, Margem do EBITDA, Imposto /Ativo Total, Rácio de Crescimento das Vendas, Rácio de Cobertura de Juros e Índice de Concentração industrial (HHI). Os resultados apresentados suportam, em parte, os resultados da análise dos coeficientes *standardized* (à exceção das variáveis Endividamento Total e Fundo de Maneio).

A última etapa da ADM é a Análise ANOVA, que significa análise de variância e é um método estatístico que procura verificar a existência de diferenças significativas entre as médias dos grupos de empresas falidas e não falidas. Esta técnica estatística será conduzida com base no *F-statistic test* sob a hipótese nula de não existência de diferenças significativas entre grupos. A tabela 16 apresenta o resultado da análise de

variância ANOVA.

**Tabela 16 - Análise ANOVA**

Variáveis	F-statistic test	P-value
Dimensão	307,4000	0,0000
Liquidez Geral	36,0750	0,0000
Liquidez Imediata	78,0110	0,0000
Fundo de Maneio	14,0020	0,0002
Inverso da Liquidez Geral	4,7816	0,0288
Solvabilidade	80,3470	0,0000
Endividamento Total	81,6600	0,0000
Rotação do Ativo	33,5380	0,0000
Endividamento de Curto Prazo	70,6720	0,0000
PM_Inventários em Armazém	12,3480	0,0004
Rácio de Eficiência	173,1500	0,0000
Rentabilidade Capitais Próprios	11,0530	0,0009
Rentabilidade Operacional do Ativo	143,4700	0,0000
Margem do EBITDA	102,1100	0,0000
Imposto /Ativo Total	116,2500	0,0000
Rácio de crescimento das vendas	123,5700	0,0000
Rácio de Cobertura de Juros	441,1400	0,0000
Idade	2,4473	0,1177
Volatilidade das vendas	0,5398	0,4625
Índice de concentração industrial (HHI)	1237,3000	0,0000

**Nota:** A Tabela 16 apresenta os resultados da análise de variância ANOVA para cada variável independente. As variáveis são as descritas no Subcapítulo 4.2.2. Os coeficientes são considerados estatisticamente significativos a um nível igual ou inferior a 5%.

O referido teste é considerado uma boa alternativa para fundamentar a remoção das variáveis sem capacidade explicativa e as variáveis são excluídas se o p-value associado ao teste for superior a 5%.

Conforme os resultados apresentados na Tabela 16, podemos concluir que, à exceção das variáveis Idade e Volatilidade das vendas, as restantes variáveis apresentam significância estatística, i.e., têm elevada capacidade na discriminação entre grupos de empresas falidas/não, falidas.

Resumindo, nesta primeira fase da análise empírica, onde conduzimos a ADM, concluímos que de acordo com a análise dos coeficientes *standardized* as variáveis que parecem ser melhor preditores da situação de falência. ou seja. que demonstram maior poder de discriminação entre os grupos, são as seguintes: Dimensão, Fundo de

Maneio, Endividamento Total, Rácio de Eficiência, Rentabilidade Operacional do Ativo, Rácio de crescimento das vendas, Rácio de Cobertura de Juros, e Índice de concentração industrial (HHI). Em relação aos resultados da análise da correlação de Pearson destacam-se as variáveis Dimensão, Rácio de Eficiência, Rentabilidade Operacional do Ativo, Rácio de crescimento das vendas, Rácio de Cobertura de Juros e Índice de concentração industrial (HHI), que corroboram os resultados da análise *standardized* e que também são consistentes com os resultados da análise ANOVA.

## 5.2 Análise Logística (*Logit Analysis*)

Tal como descrito no Capítulo 3 (Metodologia), a análise logit é a metodologia principal para a seleção das variáveis e definição do modelo final de previsão de falência. Esta decisão é baseada na evidência demonstrada na literatura (p.e., Ohlson, 1980), que considera esta metodologia como mais robusta e consistente por relaxar os pressupostos impostos pela ADM, como são disso exemplo o pressuposto da distribuição normal das variáveis independentes, a igualdade das matrizes de variância e covariância para os dois grupos (empresas falidas e empresas saudáveis) e a proporcionalidade das observações incluídas em cada grupo.

O estudo de Ohlson (1980) foi pioneiro na aplicação desta metodologia, surge com a motivação de apresentar um modelo *logit* de previsão de falência com maior capacidade preditiva do que o modelo *Z-score* de Altman (1968, 1983).

Assim, motivados pelo estudo empírico de Ohlson (1980), o nosso objetivo é propor um modelo logístico com capacidade de previsão de falência até 3 anos antes da ocorrência desse evento. Desse modo, o modelo logit especificado na equação (4) será estimado:

$$Prob (Falência = 1) = \beta' Z_{i,t-1} + \lambda_j + \gamma_t + \varepsilon_{it} \quad (4)$$

Onde a probabilidade de falência é uma *dummy* que classifica as empresas falidas com o valor de 1 ( $y=1$ ) e zero caso contrário. O vetor  $Z_{i,t-1}$  contém as variáveis independentes (que são os preditores que influenciam o comportamento da *dummy*) descritas no Subcapítulo 4.2.2 e desfasadas 1 ano.  $\lambda_j$  e  $\gamma_t$  controlam a heterogeneidade não verificada em relação à indústria (país) e aos anos; termo  $\varepsilon_{it}$  é o erro da regressão

logística.

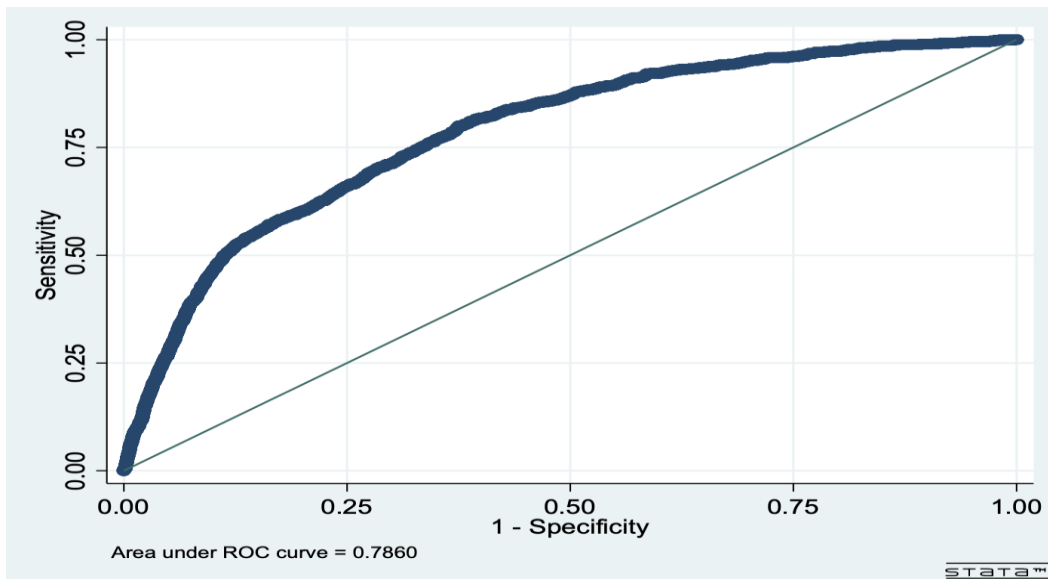
Na análise logit. a eficiência do modelo pode ser analisada mediante uma representação gráfica denominada ROC (*Receiver Operating Characteristic Curve*). A curva ROC indica o nível de eficácia do modelo com base nas seguintes medidas: *Sensitivity* indica a percentagem de casos verdadeiros positivos e *Specificity* indica a percentagem de casos verdadeiros negativos, i.e., esta análise apresenta a taxa de verdadeiros positivos *versus* a taxa de verdadeiros negativos, considerando diferentes cenários, alterando para esse efeito a variável dependente do modelo. Com base nestas duas medidas é possível encontrar o ponto de *cutoff* que minimize o risco de se incorrer em erro.

O Erro Tipo I (falsos positivos) pode ser medido como  $(1 - \text{Specificity})$ . e o Erro Tipo II (falsos negativos) pode ser medido como  $(1 - \text{Sensitivity})$ . A partir desta identificação, é possível estimar o *cutoff point* que minimize os erros.

O Gráfico 2 representa a curva ROC e a “area under ROC curve” significa a precisão com que a variável *Status* identifica de forma verdadeira as empresas falidas. Neste caso, obteve-se uma precisão de cerca de 78,60%, que é superior ao resultado de 75,05% obtido na ADM. De acordo com Hosmer, Lemeshow e Sturdivant (2013, p.177), o ponto de *breakeven* da precisão dos modelos de previsão é de 70%, o que significa que a capacidade preditiva do modelo deve superar os 70%.

Após o processo de seleção de variáveis, a curva ROC será novamente apresentada para se observarem as alterações nas mitigações do erro.

**Gráfico 2 - ROC (Receiver Operating Characteristic Curve)**



**Nota:** O Gráfico apresenta representação gráfica da ROC Curve “Area under ROC curve” representa uma medida de discriminação que avalia entre 0 a 1 a precisão com que as empresas falidas estão corretamente classificadas, tendo em conta as variáveis independentes descritas no Subcapítulo 4.2.2.

Posto isto, a próxima fase da análise logística será a *standardização* ou normalização dos coeficientes de forma a melhorar a análise comparativa da magnitude dos coeficientes e o seu poder explicativo.

A Tabela 17 apresenta resultados da estimação da equação (4) antes do processo de estandardização dos coeficientes. Os coeficientes da referida tabela são apresentados como *log odds ratios*, pelo que é necessária a sua conversão numa estimativa de probabilidade, que é efetuada de acordo com a equação (5) A título exemplificativo, o coeficiente estimado para a variável Endividamento Total é de -1.7092. Esse resultado indica que mantendo todas as outras variáveis independentes constantes (*ceteris paribus*), um aumento de uma unidade na variável Endividamento Total, implica uma redução de 1.7092 na probabilidade logarítmica da variável dependente falência. Outra interpretação, recorrendo à conversão em probabilidade através da equação definida em (5), o coeficiente da variável Endividamento Total é 0.153<sup>10</sup>, o que significa que a probabilidade de falência aumenta 15,3% quando o Endividamento Total diminui 1%, *ceteris paribus*.

<sup>10</sup> A probabilidade de falência é:  $P(Y=1) = \exp(-1,7092) / (1 + \exp(-1,7092))$ .

$$P = \frac{e^{\log-odds}}{1 + e^{\log-odds}} \quad (5)$$

**Tabela 17 - Resultados antes do Processo de Estandarização das Variáveis**

Variáveis	Coefficientes	D.P	Z-test	p-value	
Dimensão	0,0180	0,0233	0,7700	0,4410	
Liquidez Geral	-0,0069	0,0181	-0,3800	0,7050	
Liquidez Imediata	-0,1402	0,0590	-2,3800	0,0170	**
Fundo de Maneio	-0,4617	0,1448	-3,1900	0,0010	***
Inverso da Liquidez Geral	-0,1536	0,0363	-4,2300	0,0000	***
Solvabilidade	-0,4070	0,0571	-7,1300	0,0000	***
Endividamento Total	-1,7092	0,1720	-9,9400	0,0000	***
Rotação do Ativo	-0,1885	0,0323	-5,8300	0,0000	***
Endividamento de Curto Prazo	0,1080	0,1037	1,0400	0,2970	
PM_Inventários em Armazém	-0,0008	0,0001	-5,4400	0,0000	***
Rácio de Eficiência	0,5235	0,9781	0,5400	0,5920	
Rentabilidade Capitais Próprios	0,0451	0,0356	1,2700	0,2050	
Rentabilidade Operacional do Ativo	-0,8583	0,9515	-0,9000	0,3670	
Margem do EBITDA	-0,0919	0,2470	-0,3700	0,7100	
Imposto /Ativo Total	4,2450	1,9264	2,2000	0,0280	**
Rácio de crescimento das vendas	-0,9928	0,0786	-12,6300	0,0000	***
Rácio de Cobertura de Juros	-0,6224	0,0240	-25,9200	0,0000	***
Idade	-0,2301	0,0396	-5,8000	0,0000	***
Volatilidade das vendas	0,0000	0,0002	-0,1800	0,8570	
Índice de concentração industrial (HHI)	1,0659	0,0393	27,1300	0,0000	***

**Nota:** A Tabela 17 apresenta resultados da estimação da regressão (4) antes do processo de estandarização dos coeficientes. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*) e 5% (\*\*).

Por sua vez, a Tabela 18 apresenta resultados da estimação da regressão (4) depois do processo de estandarização dos coeficientes, cujo principal objetivo é mitigar a diferença de escala entre as variáveis.

Comparando os resultados da Tabela 17 e Tabela 18 (após o processo de estandarização), observa-se uma diferença significativa na magnitude do coeficiente e no seu poder explicativo. Após o processo de estandarização os coeficientes apresentam valores entre 0 e 1 o que indica que o seu poder preditivo; só as variáveis Rácio de Eficiência e Rentabilidade Operacional do Ativo excedem o valor de 1 devido o seu elevado grau de correlação, conforme anteriormente detetado na ADM.



**Tabela 18 – Resultados após o Processo de estandardização das Variáveis**

Variáveis	Coefficientes	D.P	Z-test	p-value	
Dimensão	-0,2615	0,0090	-29,1300	0,0000	***
Liquidez Geral	0,0099	0,0143	0,7000	0,4870	
Liquidez Imediata	0,1641	0,0128	12,7700	0,0000	***
Fundo de Maneio	-0,1396	0,0174	-8,0400	0,0000	***
Inverso da Liquidez Geral	0,1530	0,0139	11,0400	0,0000	***
Solvabilidade	0,2484	0,0137	18,1500	0,0000	***
Endividamento Total	0,3539	0,0185	19,1700	0,0000	***
Rotação do Ativo	0,0708	0,0103	6,9000	0,0000	***
Endividamento de Curto Prazo	-0,0695	0,0088	-7,8900	0,0000	***
PM_Inventários em Armazém	0,1126	0,0087	12,8700	0,0000	***
Rácio de Eficiência	2,6959	0,0860	31,3400	0,0000	***
Rentabilidade Capitais Próprios	-0,0147	0,0089	-1,6600	0,0960	*
Rentabilidade Operacional do Ativo	-3,5526	0,0901	-39,4500	0,0000	***
Margem do EBITDA	-0,0306	0,0145	-2,1100	0,0340	**
Imposto /Ativo Total	0,9253	0,0166	55,8800	0,0000	***
Rácio de crescimento das vendas	-0,0532	0,0073	-7,3300	0,0000	***
Rácio de Cobertura de Juros	-0,2938	0,0097	-30,2000	0,0000	***
Idade	0,0568	0,0080	7,1000	0,0000	***
Volatilidade das vendas	-0,0043	0,0118	-0,3600	0,7180	
Índice de concentração industrial (HHI)	0,0999	0,0074	13,5900	0,0000	***

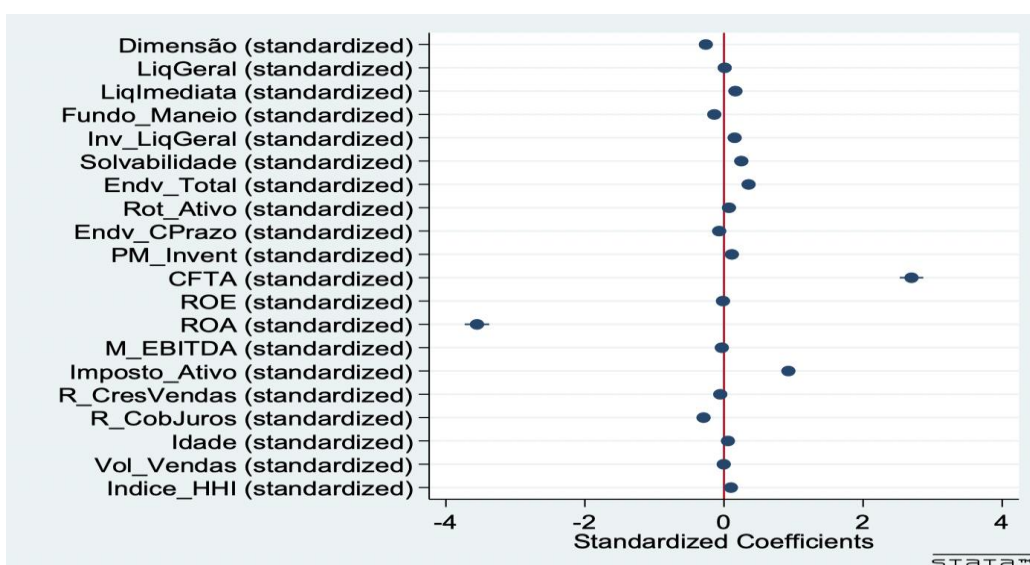
**Nota:** A Tabela 18 apresenta resultados da estimação da regressão (4) após o processo de estandardização dos coeficientes. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*) e 5% (\*\*).

De um modo geral, o processo de estandardização possibilitou a obtenção de resultados mais robustos e estatisticamente significativos, sendo que da Tabela 17 para a 18, os coeficientes das variáveis Dimensão, Endividamento de Curto Prazo, Rácio de Eficiência e Rentabilidade Operacional do Ativo ganham significância estatística ao nível de 1%, enquanto a variável Margem do EBITDA ganhou significância estatística ao nível 5%. A título de exemplo, o coeficiente da variável Rácio de Eficiência de 2,6959 indica que a variação de 1 unidade no desvio padrão na referida variável resulta num aumento de 2,6959 na variável dependente. Esta análise dos coeficientes estandardizados é acompanhada pela representação gráfica (gráfico 3) da significância dos mesmos.

A análise gráfica possibilita a observação dos coeficientes estatisticamente

diferentes de zero (i.e., os mais afastados de zero), o que significa que essas variáveis apresentam maior poder explicativo. Assim, e conforme os resultados apresentado no Gráfico 3, observamos que as variáveis Dimensão, Solvabilidade, Endividamento Total, Rácio de Eficiência, Rentabilidade Operacional do Ativo, Imposto/Ativo Total e Rácio de Cobertura de Juros apresentam coeficientes *standardized* visivelmente mais afastados de zero. confirmando a análise dos coeficientes apresentados na Tabela 18.

**Gráfico 3 – Significância dos Coeficientes *Standardized***



**Nota:** O Gráfico 3 representa a significância estatística das variáveis independentes descritas no Subcapítulo 4.2.2 após sofrerem um processo de padronização dos coeficientes.

Resumindo, os métodos utilizados até ao momento para a seleção das variáveis – ADM e análise logit -conduziram a respostas não totalmente consensuais, pelo que a seleção final será efetuada com base na significância demonstrada por cada variável ao longo das várias análises. Assim, importa salientar que as variáveis Dimensão, Endividamento Total, Rácio de eficiência, Rentabilidade Operacional do Ativo (ROA), Rácio de cobertura de juros são as variáveis comuns que se destacaram nas análises ADM e Logit.

No entanto, deste conjunto de variáveis, serão previamente escolhidas para integrarem o modelo final as variáveis Dimensão, Endividamento Total e Rácio de cobertura de juros, à qual se juntará a variável Solvabilidade, que apesar de não ter apresentado significância estatística na ADM demonstrou ser significativa na análise logística. Outra variável que deverá ser incluída no modelo final é o Índice de concentração industrial (HHI) que apresentou um elevado poder preditivo em todas

as etapas da ADM e é estatisticamente significativa na análise logística. Adicionalmente, as variáveis Rácio de Eficiência e Rentabilidade Operacional do Ativo serão integradas no modelo de forma sequencial porque, e embora tenham apresentado capacidade preditiva na ADM e análise logit, não devem ser incluídas ao mesmo tempo no modelo de regressão logística devido aos problemas de correlação entre elas reportados nas análises anteriores.

### 5.3 Propensity Score Matching

Apesar de ser considerada uma técnica que fornece resultados mais robustos do que a análise discriminante, a análise logit apresenta algumas limitações que podem afetar os resultados, como é o caso da desproporcionalidade de observações existente entre os grupos de empresas falidas e não falidas, acrescendo também as diferenças o significativas observadas nas características principais destes grupos conforme observado, por exemplo, na Tabela 11 do Subcapítulo 4.5. O método *Propensity Score Matching* (PSM), proposto por Rosenbaum e Rubin (1983), permite ultrapassar essas limitações. O propósito do PSM é fazer corresponder a cada empresa do grupo de empresas falidas – amostra de tratamento - uma (ou várias) empresas do grupo de não falidas – grupo de controlo, com o objetivo de mitigar o enviesamento resultante da desproporcionalidade da amostra, garantindo assim a consistência e a segurança dos resultados da análise, conforme descrito no Subcapítulo 3.3.

A primeira etapa para a aplicação do *Propensity Score Matching* (PSM) é a estimação do *propensity score* (pscore), ou seja, a probabilidade de uma empresa saudável entrar numa situação de falência e, para tal, os parâmetros definidos para o *matching* entre as empresas falidas/não falidas são: o ano, a indústria, o país e a variável Dimensão, *covariate* usado em estudos anteriores com base nas métricas de comparabilidade das empresas (p.e., Beaver, 1966).

O resultado da estimação do *propensity score* é apresentada na Tabela 19 e indica-nos que o coeficiente da *covariate* escolhida é estatisticamente significativa ao nível de 1%.

**Tabela 19 – Estimação do *Propensity Score***

Variável Dependente: Falência	Coefficientes	D.P	Z-test	p-value	
Dimensão	0,1354	0,0071	19,2000	<b>0,0000</b>	***
Constante	-3,0371	0,0484	-62,7900	<b>0,0000</b>	***

**Nota:** A Tabela 19 apresenta os resultados da estimação do *Propensity Score* para a variável Dimensão. Os coeficientes são estatisticamente significativos ao nível de 1%.

A segunda etapa corresponde à estimação do *Propensity Score Matching* (PSM) a partir do *propensity score*. Para tal, numa primeira fase divide-se a amostra em vários blocos com características semelhantes. De seguida, utilizamos as técnicas *Nearest Neighbour Matching* (Rosenbaum & Rubin, 1983) e *Caliper Matching with Noreplacement* (Cochran & Rubin, 1973), que representa o intervalo máximo e variação da probabilidade do *Propensity Score*. A técnica *Nearest Neighbour Matching* significa que tendo em conta o *Propensity Score* estimado, cada empresa falida será comparada a uma(várias) empresa(s) saudável(is) mais próxima(s), enquanto a aplicação do *Caliper Matching with Noreplacement* significa que cada empresa saudável será combinada apenas uma vez com uma empresa falida (matching 1:1) desde que seja possível fazer *matching* entre as observações. O *caliper* adequado corresponde a 25% do desvio padrão do *propensity score* e foi estimado segundo Cochran e Rubin (1973) e Rosenbaum e Rubin (1983).

O objetivo da estimação do PSM é de garantir a qualidade do *matching*, efetuado com base na variável Dimensão, por ano, indústria e país, entre os grupos de tratamento e controlo, analisando-se no final a existência de diferenças estatisticamente significativas entre os grupos. A Tabela 20 apresenta os resultados dessa estimação, onde é possível verificar que a qualidade do *matching* esta assegurada, i.e., não existem diferenças estatisticamente significativas para os grupos de tratamento e de controlo *Matched*, o que significa que a média da variável Dimensão entre os grupos de empresas de tratamento e controlo é semelhante após o *matching*.

**Tabela 20 – Resultados do teste realizado à qualidade do *Matching***

Variável Dependente: Falência	Grupo	Média		t-test	p-value
		Tratamento	Controlo		
Dimensão	<i>Unmatched</i>	6,898	6,342	19,690	<b>0,000</b>
	<i>Matched</i>	6,810	6,808	0,050	<b>0,960</b>

**Nota:** A Tabela 20 apresenta os resultados do teste realizado à qualidade do *matching* da variável Dimensão. Observamos as diferenças entre os grupos de Tratamento e Controlo.

Nesta fase, é importante aferir a adequação do modelo Probit que estima o *propensity score*, que é analisada através rácio de máxima verossimilhança (*Likelihood Ratio – LR*) com recurso ao teste qui-quadrado ( $\chi^2$ ).

Na Tabela 21 podemos observar através do valor do teste  $\chi^2$  (p-value de 96%) que a condição de adequação do modelo Probit é verificada.

**Tabela 21 — Análise Goodness-of-fit do Propensity Score**

Grupo	Pseudo R <sup>2</sup>	LR $\chi^2$	p-value
<i>Unmatched</i>	0,018	383,220	<b>0,000</b>
<i>Matched</i>	0,000	0,000	<b>0,960</b>

**Nota:** A Tabela 21 apresenta a qualidade do modelo probit testado por meio do rácio de máxima verossimilhança (*Likelihood Ratio LR*)  $\chi^2$ . A hipótese nula do teste é rejeitada com um nível de significância inferior a 5%, situação que não se verifica.

Em suma, os resultados da estimação do *propensity score* usando a *covariate* Dimensão revelaram que a qualidade do *matching* está assegurada.

Adicionalmente, ainda podemos efetuar uma análise univariada *Propensity Score Matching* com *Difference-in- Differences* (PSM-DID) para avaliar se após o *matching* ainda existem diferenças significativas entre grupos. Esta análise foi aplicada à variável Endividamento Total, para se analisar a diferença no grupo *matched* entre as empresas tratadas e de controlo.

**Tabela 22 — Resultados da Análise PSM-DID**

Variável	Grupo	Média		Diferenças	Desvio Padrão	t-test
		Tratamento	Controlo			
Endividamento Total	<i>Unmatched</i>	0,8536	0,7390	0,1146	0,0116	9,9000
	<i>Matched</i>	0,8524	0,6722	<b>0,1802</b>	0,0151	<b>11,9300</b>

**Nota:** A Tabela 22 apresenta os resultados da análise PSM-DID, que testa a diferença entre grupos em relação à média da variável Endividamento Total.

A Tabela 22 apresenta os resultados da análise PSM-DID, onde se pode verificar que no grupo *Matched* existe uma diferença estatisticamente significativa de 18,02 pontos percentuais entre a média do grupo de tratamento e do grupo de controlo.

## 5.4 Apresentação dos Resultados

### 5.4.1 Modelo Logit com capacidade preditiva Um, Dois e Três Anos Antes do Evento de Falência

De acordo com o objetivo definido na presente dissertação - estimar um modelo capaz de prever falência para as PME da Península Ibérica baseado em informação contabilística e financeira - nos Subcapítulos 5.1 e 5.2 foram conduzidas algumas análises preliminares que nos permitiram concluir que as variáveis Dimensão, Solvabilidade, Endividamento Total, Rácio de Cobertura de Juros e Índice de concentração industrial (HHI) são os mais significativos determinantes da situação de falência, pelo que deverão ser incluídas no modelo final. A equação (6) apresenta o modelo incluindo essas variáveis, onde a variável dependente é uma *dummy* que classifica as empresas como falidas quando o valor de 1 zero no caso contrário.

$$\begin{aligned} Prob(Falência = 1) = & \alpha + \beta_1 Dimensão + \\ & \beta_2 Solvabilidade + \beta_3 Endividamento Total + \beta_4 Rácio de Cobertura de Juros + \\ & \beta_5 Índice de Concentração Industrial(HHI) + \varepsilon_{it} \end{aligned} \quad (6)$$

A Tabela 23 apresenta os resultados da estimação da regressão (6), para 1, 2 e 3 anos antes da falência. Analisando os respetivos resultados, podemos verificar em relação ao teste de significância global dos coeficientes Chi<sup>2</sup>-test que as variáveis independentes incluídas no modelo apresentam uma elevada capacidade para prever eventos de falência para um nível de significância de 1% até 3 anos antes do evento acontecer, i. e., o conjunto de variáveis incluídas no modelo tem capacidade para explicar o comportamento da variável dependente.

**Tabela 23 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência**

Variável dependente : Falência Modelo	1 Ano Antes			2 Ano Antes			3 Ano Antes		
	Coef,	P-value	z	Coef,	P-value	z	Coef,	P-value	z
Dimensão	0,3571	0,0000	***	0,3431	0,0000	***	0,3208	0,0000	***
Solvabilidade	-0,1551	0,0000	***	-0,1542	0,0000	***	-0,1585	0,0000	***
Endividamento Total	0,1472	0,0140	**	0,0791	0,2500		0,0526	0,4910	
Rácio de Cobertura de Juros	-0,3156	0,0000	***	-0,3066	0,0000	***	-0,3032	0,0000	***
Índice de concentração industrial (HHI)	1,4883	0,0000	***	1,4918	0,0000	***	1,4965	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000			0,0000			0,0000		
<b>Pseudo R2</b>	0,1410			0,1343			0,1279		
<b>Cutoff point</b>	0,0189			0,0170			0,0145		
<b>% Obs, Corretamente Classificadas</b>	75,24%			76,02%			76,41%		
<b>Curva ROC</b>	81,97%			81,71%			81,55%		
<b>N.º Observações</b>	112.827			112.559			112.298		

**Nota:** A Tabela 23 apresenta os resultados da estimação da equação (6), nomeadamente os coeficientes e o respetivo o *p*-value. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis independentes são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*) e 5% (\*\*).

À exceção da variável Endividamento Total, que apresenta um coeficiente estatisticamente significativo ao nível de 5% no modelo antes 1 ano da falência, perdendo a significância nos modelos seguintes, todas as restantes variáveis apresentaram coeficientes estatisticamente significativos ao nível de 1%. Não obstante da importância que cada coeficiente apresenta individualmente, é de realçar que é o conjunto dos coeficientes que possibilita determinar a probabilidade de uma empresa entrar em situação de falência. Deste modo, podemos observar que o modelo apresentado é eficiente dado que, em média, classifica corretamente mais 75% das observações nos 3 modelos. Em relação à capacidade de discriminação entre empresas saudáveis e falidas, apresenta um nível de precisão médio de 81% (curva ROC), o que é considerado um nível de precisão elevado de acordo com Hosmer, Lemeshow & Sturdivant (2013, p.177). Observa-se ainda uma ligeira diminuição da capacidade do modelo à medida que o desfasamento temporal aumenta em relação à declaração do estado de falência o que é consistente com o estudo de Ohlson (1980).

Tendo em conta os resultados acima obtidos, é relevante efetuar um teste à qualidade da inferência estatística da regressão estimada. Para tal, a equação (6) foi estimada considerando a opção de erro-padrão *White Heteroskedasticity-Adjusted*



*Robust* para correção de problemas de autocorrelação e heteroscedasticidade. Na Tabela 24 podemos verificar os resultados da estimação dos coeficientes com a aplicação da opção de controlo dos erros-padrão White-robust (Cenário A) e com opção de cluster dos erros-padrão por indústria (cenário B).

**Tabela 24 — Resultados da estimação do modelo Logit com os Erros-Padrão Robust e por Cluster**

Cenário A – Estimação com Erros-Padrão Robustos

Variável Dependente: Falência	Coefficientes	D.P	Z-test	P-value	
Dimensão	0,3843	0,0216	17,7900	0,0000	***
Solvabilidade	-0,1574	0,0247	-6,3400	0,0000	***
Endividamento Total	0,2391	0,0541	4,4200	0,0000	***
Rácio de Cobertura de Juros	-0,3102	0,0183	-16,9600	0,0000	***
Índice de concentração industrial (HHI)	1,3407	0,0483	27,7400	0,0000	***
Constante	-7,8540	0,2284	-34,3900	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000				
<b>Pseudo R2</b>	0,1342				
<b>Cutoff point</b>	0,0191				
<b>% Obs. Corretamente Classificadas</b>	73,68%				
<b>Curva ROC</b>	81,05%				
<b>N.º Observações</b>	112.957				

Cenário B – Estimação com Erros-Padrão Robustos por Cluster

Variável Dependente: Falência	Coefficientes	D.P	Z-test	P-value	
Dimensão	0,3843	0,0432	8,9000	0,0000	***
Solvabilidade	-0,1574	0,0560	-2,8100	0,0050	***
Endividamento Total	0,2391	0,1055	2,2700	0,0230	**
Rácio de Cobertura de Juros	-0,3102	0,0440	-7,0500	0,0000	***
Índice de concentração industrial (HHI)	1,3407	0,0951	14,0900	0,0000	***
Constante	-7,8540	0,4300	-18,2600	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,000				
<b>Pseudo R2</b>	0,134				
<b>Cutoff point</b>	0,019				
<b>% Obs. Corretamente Classificadas</b>	73,68%				
<b>Curva ROC</b>	81,05%				
<b>N.º Observações</b>	112.957				

**Nota:** A Tabela 24 apresenta os resultados da estimação da equação (6) com os erros-padrão Robust (Cenário A) e por Cluster (Cenário B). A Variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*), 5% (\*\*) e 10% (\*).

Os resultados apresentados na Tabela 24 são consistentes com os resultados anteriores (Tabela 23), no que respeita à magnitude e significância dos coeficientes. De salientar ainda que a significância do coeficiente da variável Endividamento Total, de 5% no cenário A e 1% no cenário B.

O nível de precisão do modelo (curva ROC) manteve-se em cerca 81%, e a percentagem de observações corretamente classificadas é de aproximadamente 74%, ligeiramente inferior à percentagem apresentada na Tabela 23.

Com o objetivo de avaliar a robustez do modelo estimado, a equação (6) foi reestimada com a inclusão das variáveis Rácio de Eficiência (Cenário A) e Rentabilidade Operacional do Ativo (Cenário B), que demonstraram capacidade preditiva na ADM e análise logística, mas devido ao seu elevado grau de correlação não foram incluídas em simultâneo no modelo final. A tabela 25 apresenta os resultados. com desfasamento de 1, 2 e 3 anos antes da falência.

**Tabela 25 — Resultados da estimação do modelo Logit 1. 2 e 3 Anos Antes da Falência**

Cenário A – Inclusão da Variável Rácio de Eficiência

Variável dependente : Falência Modelo	1 Ano Antes			2 Ano Antes			3 Ano Antes		
	Coef.	P-value	z	Coef.	P-value	z	Coef.	P-value	z
Dimensão	0,3593	0,0000	***	0,3432	0,0000	***	0,3207	0,0000	***
Solvabilidade	-0,1644	0,0000	***	-0,1544	0,0000	***	-0,1582	0,0000	***
Endividamento Total	0,0785	0,2400		0,0776	0,2860		0,0546	0,4860	
Rácio de Cobertura de Juros	-0,2975	0,0000	***	-0,3062	0,0000	***	-0,3038	0,0000	***
Índice de concentração industrial (HHI)	1,4856	0,0000	***	1,4918	0,0000	***	1,4966	0,0000	***
<b>Rácio de Eficiência</b>	<b>-0,4735</b>	<b>0,0300</b>	<b>**</b>	<b>-0,0105</b>	<b>0,9660</b>		<b>0,0144</b>	<b>0,9550</b>	
Constante	-8,0804	0,0000	***	-8,1024	0,0000	***	-8,1421	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000			0,0000			0,0000		
<b>Pseudo R2</b>	0,1412			0,1343			0,1279		
<b>Cutoff point</b>	0,0187			0,0170			0,0144		
<b>% Obs. Corretamente Classificadas</b>	75,09%			76,07%			76,36%		
<b>Curva ROC</b>	81,99%			81,71%			81,55%		
<b>N.º Observações</b>	112.826			112.558			112.297		

Cenário B – Inclusão da Variável Rentabilidade Operacional do Ativo

Variável dependente : Falência Modelo	1 Ano Antes			2 Ano Antes			3 Ano Antes		
	Coef.	P-value	z	Coef.	P-value	z	Coef.	P-value	z
Dimensão	0,3576	0,0000	***	0,3429	0,0000	***	0,3205	0,0000	***
Solvabilidade	-0,1598	0,0000	**	-0,1497	0,0000	***	-0,1524	0,0000	***
Endividamento Total	0,1151	0,0770	*	0,1107	0,1190		0,0948	0,2150	
Rácio de Cobertura de Juros	-0,3067	0,0000	***	-0,3159	0,0000	***	-0,3158	0,0000	***
Índice de concentração industrial (HHI)	1,4869	0,0000	***	1,4930	0,0000	***	1,4980	0,0000	***
<b>Rentabilidade Operacional do Ativo</b>	<b>-0,2201</b>	<b>0,2920</b>		<b>0,2301</b>	<b>0,3190</b>		<b>0,3083</b>	<b>0,2100</b>	
Constante	-8,0462	0,0000	***	-8,0810	0,0000	***	-8,1161	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000			0,0000			0,0000		
<b>Pseudo R2</b>	0,1411			0,1343			0,1280		
<b>Cutoff point</b>	0,0179			0,0165			0,0146		
<b>% Obs. Corretamente Classificadas</b>	74,13%			75,41%			76,55%		
<b>Curva ROC</b>	81,98%			81,71%			81,54%		
<b>N.º Observações</b>	112.827			112.559			112.298		

**Nota:** A Tabela 25 apresenta os resultados da estimação da equação (6) com os erros-padrão Robust (Cenário A) e por Cluster (Cenário B). A Variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*), 5% (\*\*) e 10% (\*).

Decorre da análise dos resultados que o coeficiente da variável Rácio de Eficiência é apenas estatisticamente significativo ao nível de 5% para o modelo 1 ano antes da situação de falência, enquanto o coeficiente da variável Rentabilidade Operacional do Ativo não apresentou qualquer significância estatística para os 3 anos antes da falência. De referir que a variável Endividamento Total perde significância estatística no cenário B e não apresenta qualquer significância estatística no cenário A. A capacidade preditiva média para os 2 cenários mantém-se acima de 81%, apresentando uma percentagem de acerto de observações corretamente classificadas de entre 74% e 76%, o que corrobora os resultados anteriormente apresentados nas Tabelas 23 e 24.

Em resumo, a evidência demonstra que o conjunto de variáveis selecionadas nos dois cenários (com a inclusão da variável Rácio de Eficiência ou Rentabilidade operacional do Ativo) possui capacidade de explicar o evento de falência conforme teste de significância global dos coeficientes Chi2 ( $p$ -value inferior a 1%).

De forma a testar-se a robustez do modelo definido na equação (6), foi incluído um conjunto de *dummies* para controlar para efeitos não observáveis por ano e por indústria (conforme classificadas na Tabela A do Apêndice) e por País e até 3 anos antes da falência. Deste modo, pretende-se analisar se, controlando-se para os efeitos temporais e características não observáveis por indústria/país, os novos resultados serão substancialmente diferentes dos anteriores. A Tabela 26 apresenta os resultados.

**Tabela 26 — Resultados da estimação do modelo Logit 1. 2 e 3 Anos Antes da Falência. controlando para Efeitos Temporais**

Variável dependente: Falência Modelo	1 Ano Antes			2 Ano Antes			3 Ano Antes		
	Coef.	P-value	z	Coef.	P-value	z	Coef.	P-value	z
Dimensão	0,3485	0,0000	***	0,3547	0,0000	***	0,3542	0,0000	***
Solvabilidade	-0,1353	0,0000	***	-0,1294	0,0000	***	-0,1318	0,0000	***
Endividamento Total	0,1363	0,0280	**	0,0972	0,1740		0,0997	0,2140	
Rácio de Cobertura de Juros	-0,3071	0,0000	***	-0,2705	0,0000	***	-0,2321	0,0000	***
Índice de concentração industrial (HHI)	1,5526	0,0000	***	1,5624	0,0000	***	1,5733	0,0000	***
Constante	-9,4023	0,0000	***	-9,7149	0,0000	***	-9,9933	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000			0,0000			0,0000		
<b>Pseudo R2</b>	0,1678			0,1642			0,1613		
<b>Cutoff point</b>	0,0212			0,0214			0,0208		
<b>% Obs. Corretamente Classificadas</b>	76,72%			76,70%			76,13%		
<b>Curva ROC</b>	83,61%			83,35%			83,20%		
<b>N.º Observações</b>	100.675			86.319			71.942		

**Nota:** A Tabela 26 apresenta os resultados da estimação da equação (6), controlando-se para os efeitos não observáveis (ano, indústria e país), nomeadamente, os coeficientes e o respetivo *p*-value. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*), 5% (\*\*) e 10% (\*).

Os resultados apresentados na Tabela 26 conduzem à conclusão de que a inclusão do conjunto de *dummies* por ano/indústria/país aumenta a capacidade preditiva do modelo que passou a ser, em média, de 83% (versus 81% de média nas análises anteriores) para os 3 anos antes, observando-se assim um incremento da qualidade de ajustamento dos modelos estimados para 1, 2 e 3 anos antes da falência. Adicionalmente, o modelo classifica mais de 76% das observações corretamente, o que também reflete um incremento face à percentagem de acerto apresentada na Tabela 23.

A análise seguinte apresenta a reestimação da equação (6) incluindo somente as observações *matched* obtidas durante o processo de *Propensity Score Matching*. A Tabela 27 apresenta os resultados.

**Tabela 27 — Resultados da estimação do modelo Logit 1, 2 e 3 Anos Antes da Falência: Observações *matched***

Variável Dependente: Falência Modelo	1 Ano Antes			2 Anos Antes			3 Anos Antes		
	Coef.	P-value	z	Coef.	P-value	z	Coef.	P-value	z
Dimensão	0,0182	0,6230		0,0025	0,9500		-0,0060	0,8880	
Solvabilidade	-0,1720	0,0000	***	-0,1606	0,0000	***	-0,1620	0,0000	***
Endividamento Total	0,2335	0,1010		0,1595	0,2850		0,1669	0,3130	
Rácio de Cobertura de Juros	-0,3008	0,0000	***	-0,2802	0,0000	***	-0,2436	0,0000	***
Índice de Concentração Industrial (HHI)	1,4267	0,0000	***	1,4293	0,0000	***	1,4511	0,0000	***
Constante	-1,6143	0,0180	**	-1,5309	0,0450	**	-1,7049	0,0300	**
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000			0,0000			0,0000		
<b>Pseudo R2</b>	0,2274			0,2197			0,2189		
<b>% Obs. Corretamente classificadas</b>	75,61%			75,25%			75,42%		
<b>Cutoff</b>	0,5044			0,5259			0,5151		
<b>LROC</b>	81,17%			80,70%			80,64%		
<b>Nº Observações</b>	3.370			2.885			2.413		

**Nota:** A Tabela 27 apresenta os resultados da estimação da equação (6), controlando-se para os efeitos não observáveis (ano, indústria e país) e usando as observações *matched* resultantes da aplicação técnica PSM, nomeadamente, os coeficientes e o respetivo *p*-value. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (\*\*\*), 5% (\*\*) e 10% (\*).

Observando os resultados apresentados na Tabela 27, verifica-se que à exceção dos coeficientes das variáveis Dimensão e Endividamento Total que não revelaram qualquer significância estatística, as restantes variáveis apresentaram coeficientes estatisticamente significativos ao nível de 1%. Consistente com os resultados anteriores, o teste de significância global dos coeficientes (Chi2) suporta evidência que o conjunto de variáveis explica a situação de falência. A capacidade preditiva do modelo mantém-se, em média, nos 81%, o que é consistente com os resultados das Tabelas 23-25.

#### 5.4.2 Modelo Final de Previsão de Falência

Na sequência das análises anteriores, que suportam evidência da robustez do nosso modelo de previsão de falência, a equação (6) foi reestimada para se identificarem os coeficientes que definirão o modelo final de previsão de falência, controlando para os efeitos não observáveis por ano, indústria e país. A Tabela 28 apresenta os resultados da estimação do modelo final, seguida pela respetiva curva

ROC.

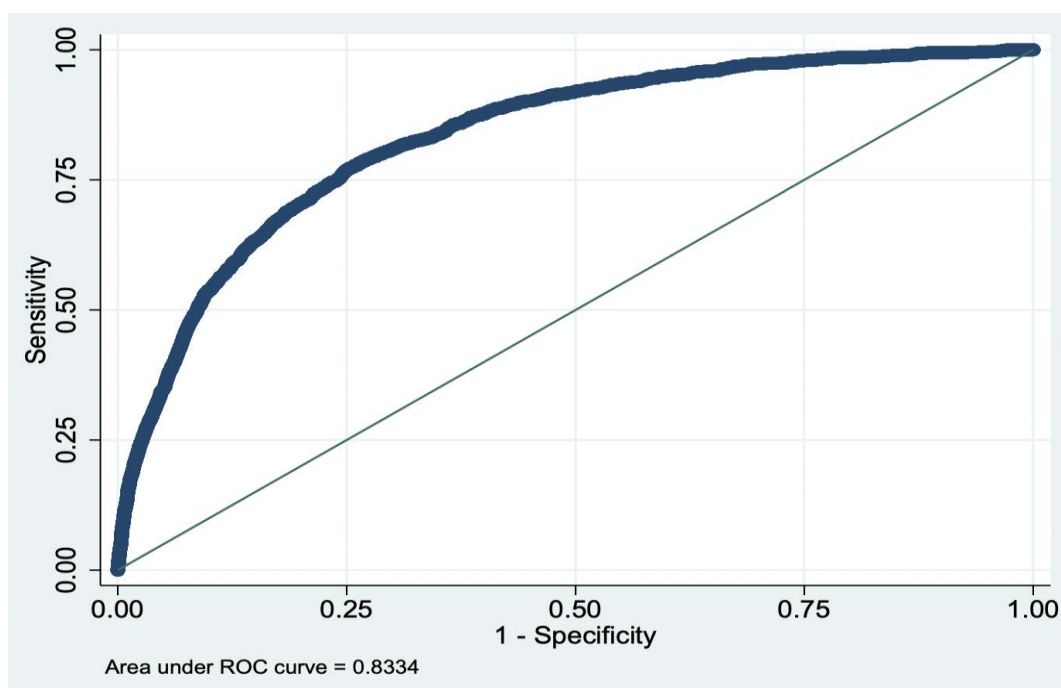
Os resultados apresentados na Tabela 28 são idênticos aos obtidos nas análises anteriores. De salientar que este modelo apresenta nível de precisão de aproximadamente de 83% (conforme apresentado no Gráfico 4) e que 75% das observações estão corretamente classificadas e um De acordo com o ponto de *break even* divulgado na literatura anterior (ver. p.e.. Hosmer. Lemeshow & Sturdivant. 2013. p.177) que ronda os 70%, o nível de precisão do nosso modelo é considerado bastante satisfatório.

**Tabela 28 — Resultados da estimação do modelo Logit Final de Previsão de Falência**

Variável Dependente: Falência	Coefficientes	D.P	Z-test	P-value	
Dimensão	0,3681	0,0224	16,4600	0,0000	***
Solvabilidade	-0,1424	0,0246	-5,7900	0,0000	***
Endividamento Total	0,2228	0,0556	4,100	0,0000	***
Rácio de Cobertura de Juros	-0,3112	0,0190	-16,3400	0,0000	***
Índice de concentração industrial (HHI)	1,4097	0,0494	28,5200	0,0000	***
Constante	-9,9469	0,3600	-27,6300	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000				
<b>Pseudo R2</b>	0,1653				
<b>Cutoff point</b>	0,0184				
<b>% Obs, Corretamente Classificadas</b>	75,06%				
<b>Curva ROC</b>	83,34%				
<b>N.º Observações</b>	112.957				

**Nota:** A Tabela 28 apresenta os resultados da estimação da equação (6, controlando para os efeitos não observáveis (ano, indústria e país). A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de 1% (\*\*\*).

**Gráfico 4 – ROC (Receiver Operating Characteristic Curve): Modelo Logit Final**



**Nota:** O Gráfico 4 representa a curva ROC da estimação da equação (6). A “Area under ROC curve” representa uma medida de discriminação que avalia entre 0 e 1 a precisão com que as empresas falidas estão corretamente classificadas. As variáveis explicativas são as descritas no Subcapítulo 4.2.2.

De acordo com os resultados apresentados na Tabela 28, o modelo geral final de previsão de falência é o seguinte:

$$\begin{aligned} \text{Score} = & -9,9469 + 0,3681 \text{ Dimensão} - 0,1424 \text{ Solvabilidade} + \\ & 0,2228 \text{ Endividamento Total} - 0,3112 \text{ Rácio de Cobertura de Juros} + \\ & 1,4097 \text{ Índice de Concentração Industrial(HHI)} \end{aligned} \quad (7)$$

A probabilidade do *score*, ou seja, de a empresa entrar em situação de falência é representada pela seguinte expressão:

$$\text{Prob}(\text{Falência} = 1) = \frac{1}{1 + e^{-\text{score}}} \quad (8)$$

O modelo estimado apresenta um *cutoff point* de 0,0184. o que significa que se o *score* for superior ao *cutoff point* a empresa será classificada com falida, caso contrário, será classificada como saudável.

O Apêndice contempla a aplicação prática do modelo (7) a uma empresa da base



de dados usada neste trabalho.

### 5.4.3 Teste de Robustez

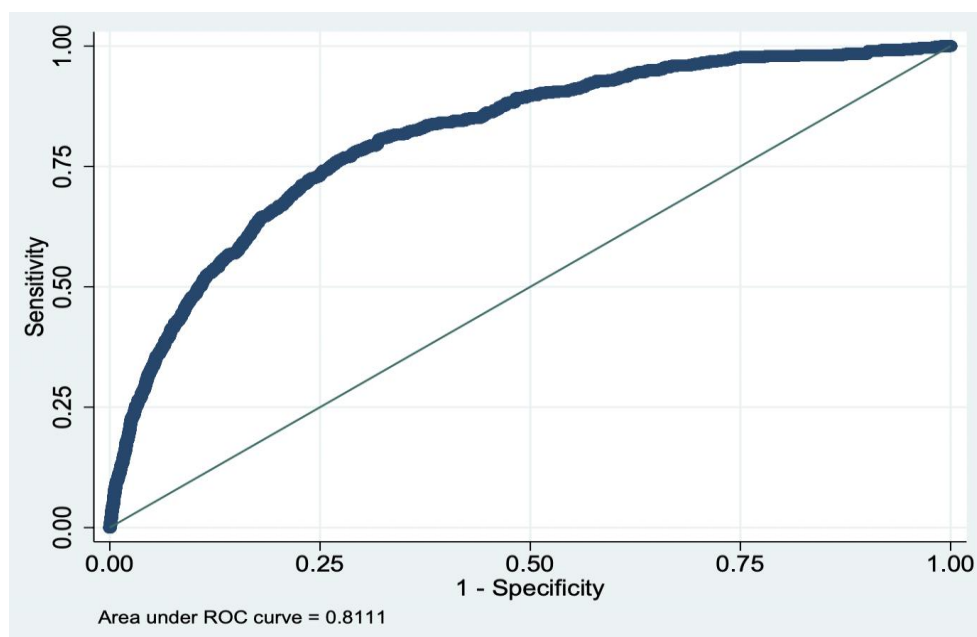
Para finalizar a análise empírica e validar a consistência dos resultados anteriores, será aplicado um teste *out-of-sample*. O objetivo desta análise é verificar a eficácia preditiva do modelo quando aplicado a uma parte temporal da nossa amostra. Para tal, estimou-se o modelo definido na equação (6) para o conjunto de observações da subamostra no período compreendido entre os anos 2012 e 2016. Com base nessa estimativa, efetuou-se uma previsão o restante período da amostra, compreendido entre os anos 2017-2020. A Tabela 29 apresenta os resultados e o Gráfico 5 apresenta a respectiva curva ROC.

**Tabela 29 — Resultados da estimação do teste Out-of-sample**

Variável Dependente: Falência	Coefficientes	D.P	Z-test	P-value	
Dimensão	0,3929	0,0280	14,0500	0,0000	***
Solvabilidade	-0,1455	0,0308	-4,7300	0,0000	***
Endividamento Total	0,1671	0,0849	1,9700	0,0490	**
Rácio de Cobertura de Juros	-0,1827	0,0243	-7,5100	0,0000	***
Índice de concentração industrial (HHI)	1,4176	0,0675	20,9900	0,0000	***
Constante	-8,8355	0,3172	-27,8600	0,0000	***
<b>Significância Global Coeficientes: Chi2-test</b>	0,0000				
<b>Pseudo R2</b>	0,127				
<b>Cutoff</b>	0,0231				
<b>% Obs, Corretamente Classificadas</b>	75,12%				
<b>ROC</b>	81,11%				
<b>Nº Observações</b>	55.424				

**Nota:** A Tabela 29 apresenta os resultados da estimação da equação (6) para uma subamostra compreendida entre 2012 e 2016. A variável dependente é uma *dummy* que classifica as empresas com 1 se falida e 0 caso contrário. As variáveis explicativas são as descritas no Subcapítulo 4.2.2. Os coeficientes são estatisticamente significativos ao nível de 1% (\*\*\*) e 5% (\*\*).

**Gráfico 5 – Curva ROC do teste *Out-of-sample***



**Nota:** O Gráfico 5 representa curva ROC da estimação da equação (6), para a Subamostra compreendida entre 2012-2016. As variáveis explicativas são as descritas no Subcapítulo 5.2.

Os resultados obtidos com o teste *Out-of-sample* apresentados na Tabela 29 são consistentes com os resultados apresentados nas análises anteriores. O modelo classifica corretamente cerca de 75% da amostra e analisando a representação gráfica do modelo, i.e., a curva ROC (Gráfico 5), constatamos que o modelo apresenta uma percentagem de precisão de cerca de 81% que suporta o elevado poder preditivo do modelo estimado.

Os resultados da Tabela 29 suportam a hipótese de que o modelo foi bem definido, tendo a capacidade preditiva para detetar se uma empresa se encontra na iminência de falir ou não.

## 6. Conclusão

Os modelos de previsão de falência empresarial têm motivado várias investigações ao longo das últimas décadas, envolvendo diferentes perspetivas económicas e diferentes países. Tanto que após as investigações pioneiras de Beaver (1966), Altman (1968) e Ohlson (1980), muitos outros autores propuseram modelos alternativos.

Tendo em conta o contexto de incerteza económica que o mundo vive devido à pandemia global causada pela COVID-19, é esperado um aumento significativo do número de falências, principalmente PME que estão mais expostas a várias fontes de risco. Esse facto motivou a seleção deste segmento empresarial para o desenvolvimento e análise de uma modelo de previsão de falência. Um outro motivo reside no papel fundamental das PME na economia de qualquer país, ao promoverem o crescimento económico, gerando empregos e contribuindo para mais de metade do valor acrescentado total criado pelas empresas na EU ( European Commission, 2021).

Neste contexto, e com recurso à utilização de um conjunto de rácios financeiros, o presente estudo centrou-se em identificar as variáveis consideradas determinantes de falência de PME espanholas e portuguesas de vários setores de atividade. Para o efeito, foram selecionados e analisados 20 rácios, construídos com base na literatura anterior. A informação financeira necessária foi recolhida da base de dados ORBIS Europe. A amostra final é constituída por 16.536 PME, para um período compreendido entre os anos 2012 a 2020.

Aplicando como metodologia base a regressão logística de acordo com Ohlson (1980), vários passos foram percorridos para o desenvolvimento de um modelo final que demonstrasse elevada capacidade para a deteção de falência. Na primeira etapa foram selecionadas as variáveis com maior capacidade para prever falência mediante a aplicação da Análise Discriminante Múltipla (ADM) e também com recurso à Análise Logística. Esse processo resultou na seleção das seguintes variáveis: Dimensão, Solvabilidade, Endividamento Total, Rácio de Cobertura de Juros e Índice de concentração industrial (HHI) para integrarem o modelo final. O modelo final foi sujeito a diferentes análises de robustez, tais como: usada i) a metodologia *Propensity Score Matching* para mitigar problemas relacionados com a desproporcionalidade da amostra em termos de observações das empresas falidas vs não falidas. ; ii) o modelo

final foi analisado 1, 2 e 3 anos antes do evento de falência; iii) na maioria das regressões, foram incluídas *dummies* para se controlarem efeitos não observáveis ao longo do tempo, por indústria e por país; e iv) foi efetuado uma análise *out-of-sample* com o objetivo de se avaliar a capacidade preditiva do modelo quando aplicado num momento futuro.

Os resultados sugerem que o modelo final apresenta elevada precisão na deteção de falência até 3 anos antes do evento; o nível de precisão do modelo final é cerca de 83%, o que é consistente com a maioria de trabalhos cujo foco é a previsão de falência empresarial.

O modelo proposto neste trabalho é um contributo para a propagação do conhecimento empírico na matéria das metodologias aplicadas à previsão de falência empresarial, principalmente pelo facto de analisar em específico o segmento das PME, que é sempre preterido em detrimento das grandes empresas.

Importa também mencionar as limitações deste trabalho, sendo a primeira das quais a falta de informação disponibilizada pela ORBIS Europe relativamente a empresas falidas, pelo que acreditamos ser um número superior ao considerado. Dada esta limitação, parece pertinente replicar este mesmo estudo num futuro próximo incluindo as empresas classificadas como *zombies*, i.e., empresas com custos de financiamento superiores aos resultados operacionais nos últimos 3 anos como alternativa à classificação de empresas em situação de falência. Além disso, a inclusão de novas variáveis, tais como: género e competência dos gestores e/ou qualificações dos trabalhadores poderão revelar-se muito importantes na deteção da situação de falência.

## Referências Bibliográficas

Altman, E. I. (1968). Financial Ratios. Discriminant Analysis and The Prediction of Corporate Bankruptcy. *The Journal of Finance*. 23(1). 589–609. <https://doi.org/10.1111/j.1540-6261.1974.tb00057.x>.

Altman, E. I. (1983). Multidimensional graphics and bankruptcy prediction: a comment. *Journal of Accounting Research*. 297-299.

Altman, E. I., & Hotchkiss, E. (1993). *Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy*. Analyze and Invest in Distressed Debt. John Wiley & Sons.

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress Prediction in an International context: A review and empirical analysis of Altman's Z-Score model. *Journal of International Financial Management & Accounting*. 28 (2). 131-171. <https://doi.org/10.1111/jifm.12053>.

Altman, E. I. & Sabato, G. (2007). Modelling Credit Risk for SMEs: Evidence from the U.S. Market. *Abacus*. 43(3). 332–357. <https://doi.org/10.1111/j.1467-6281.2007.00234.x>.

Altman, E. I., Sabato, G., & Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk*. 6(2). 95-127.

Arslan, O., & Karan, M. B. (2009). Credit risks and internationalization of SMEs. *Journal of Business Economics and Management*. 10(4). 361-368. <https://doi.org/10.3846/1611-1699.2009.10.361-368>.

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*. 4. 71– 111. <https://doi.org/10.2307/2490171>.

Blanco-Oliver, A. J., Diéguez, A. I. I., & Alfonso. M. D. O. (2012). The Prediction of Bankruptcy of Small Firms in the UK using Logistic Regression. *Análisis financiero*. (118). 32-40.

Blum, M. (1974). Failing company discriminant analysis. *Journal of accounting research*. 1-25.

Bellovary, J., Giacomino, D., & Akers. M. (2007). A Review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education*. 33. 1-42.

Cherobim, A., Junior, A. & Rigo, C. (2010). *Administração financeira: princípios, fundamentos e práticas brasileiras*. 4.ª Edição. Rio de Janeiro (Campus): Elsevier:

Cochran, W. G. and Rubin, D. B. (1973) Controlling bias in observational studies: A review. *Sankya*. Series A. 35. 417–446.

Cook, R. A., & Nelson, J. L. (1998). A conspectus of business failure forecasting. *The Journal of finance*. 23(4). 589-609.

Charalambakis, E. J., & Garrett. I. (2019). On corporate financial distress

prediction: What can we learn from private firms in a developing economy? Evidence from Greece. *Rev Quant Finan Acc* (2019) 52:467–491. <https://doi.org/10.1007/s11156-018-0716-7>.

Dănilă, O. M. (2012). Credit Risk Assessment under Basel Accords. *Theoretical & Applied Economics*. 19(3). 77-90.

Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of accounting research*. 167-179.

Domenichelli, O. (2018). Performance, financing decisions and corporate governance of Italian medium and large private family firms. *African Journal of Business Management*. 12(19). 574-585. <https://doi.org/10.5897/AJBM2018.866>.

Du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*. 73(10-12). 2047-2060.

European Commission (2021). “2021 SME COUNTRY FACT SHEET”. Obtido de <https://ec.europa.eu/docsroom/documents/46060>.

European Central Bank (ECB). Survey on the access to finance of enterprises (SAFE) - *analytical report* (2013). Obtido de <https://ec.europa.eu/docsroom/documents/7864/attachments/1/translations>.

Frank, R., & Urbancic, D. B. A. (2005). *The power of cash flow ratios*. Department of Accounting. Mitchell College of Business University of South Alabama. Alabama.

Franzblau, A. (1958). *A primer of statistics for non-statisticians*. New York: Harcourt, Brace & World, Inc.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (2003). *Applied Statistics for the Behavioral Sciences*. 5th ed. Boston: Houghton Mifflin.

Hosmer, Jr. D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons. Hoboken, New Jersey.

Jackendoff, N. (1962). *A study of published industry financial and operating ratios*. Philadelphia: Temple University. Bureau of Economic and Business Research.

Gupta, J., Wilson, N., Gregoriou, A., & Healy, J. (2014). The effect of internationalization on modelling credit risk for SMEs: Evidence from UK market. *Journal of International Financial Markets Institutions & Money*. 31. 397-413.

Lacerda, A., & Moro, R. A. (2008). *Analysis of the predictors of default for portuguese firms* (No. w200822).

Laffarga, J. (1999). Los modelos de predicción de la insolvencia empresarial: limitaciones y utilidades. *Boletín AECA - Revista de la Asociación Española de Contabilidad y Administración de Empresas*. 48. 31-34.

MarôCo, J. (2003). *Análise estatística – com utilização do SPSS* (2.<sup>a</sup> ed.). Lisboa: Edições Sílabo.

Merwin, C. L. (1942). Financing Small Corporations in Five Manufacturing

Industries. New York: *National Bureau of Economic Research*. 1926–1936.

Morgado, A. (1998). Contribuição da análise discriminante na previsão do risco de insolvência financeira. *VII Jornadas de Contabilidade e Auditoria: "Século XXI: os novos contextos da globalização. contabilidade e auditoria"*. Coimbra: Instituto Superior de Contabilidade e Administração de Coimbra.

Neves, J.C. (1997). Diagnóstico e plano de acção para recuperação - Uma visão de mercado e de crítica ao CPEREF. *Seminário sobre Gestão e Liquidação Judicial: O Estado da Arte em Portugal*. Associação portuguesa de Gestores e Liquidatários Judiciais.

Neves, J., & Silva, J. (1998). *Modelos de risco de incumprimento à Segurança Social*. Lisboa: ISEG - Fundação para a Ciência e Tecnologia.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*. 18(1). 109–131. <https://doi.org/10.2307/2490395>.

Pereira, J. M. C. (2021). *Determinantes da insolvência de PME do norte e centro de Portugal*. Dissertação de Mestrado. Universidade do Minho.

Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*. 70(1). p. 41–55. <https://doi.org/10.1093/biomet/70.1.41>.

SAGE, Pub. (2015). Extension chapters on advanced techniques. Obtido de DiscriminantAnalysis:<https://webcache.googleusercontent.com/search?q=cache:rWq9ynQWKnEJ:https://www.researchgate.net/file.PostFileLoader.html%3Fid%3D548f502ed11b8b793c8b45bc%26assetKey%3DAS%253A273652862>.

Sera, R. (2020). *Financial Distress Prediction for Portuguese Smes*. Dissertação de Mestrado. Universidade do Minho, Portugal.

Silva, J. C. C. B. D. (2021). *Determinantes da insolvência financeira das PME em Portugal: evidência para a área metropolitana de Lisboa, Alentejo e Algarve*. Dissertação de Mestrado. Universidade do Minho, Portugal.

Zmijweski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*. 22. 59–82.

## Apêndice

**Tabela A – Definição das indústrias utilizadas no estudo de acordo com o CAE**

Indústria	Descrição	CAE
1	Agricultura, produção animal, caça, floresta e pesca	1; 2; 3
2	Indústrias Extrativas e transformadoras	7; 8; 10; 11; 13; 14; 15; 16; 17; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28; 29; 30; 31; 32; 33
3	Distribuição de água, saneamento e gestão de resíduos	37; 38; 39
4	Construção	41; 42; 43
5	Comércio por grosso, retalho e serviços de reparação	45; 46; 47
6	Serviços de transportes	49; 50; 51; 52; 53
7	Hoteis, restauração e similares	55; 56
8	Serviços de informação e de comunicação	58; 59; 60; 61; 62; 63
9	Atividades de consultoria, científicas e técnicas	69; 70; 71; 72; 73; 74; 75
10	Actividades administrativas e dos serviços de apoio	77; 78; 79; 80; 81; 82
11	Educação	85
12	Actividades de saúde e apoio social	86; 87; 88
13	Actividades associativas e recreativas	90; 91; 92; 93
14	Outras atividades e serviços	94; 95; 96; 99

**Nota:** Esta Tabela apresenta os CAE utilizados no processo de recolha da nossa amostra e definição da indústria.

### Aplicação Prática do Modelo

Decidiu-se aplicar o modelo geral final definido na equação (7) de previsão de falência a um caso particular, cuja informação foi recolhida de forma aleatória da nossa base de dados.



Assim, no ano de 2016, a empresa com nome fictício “ONDAH” apresentava os seguintes coeficientes:

- Dimensão: 4,9997;
- Solvabilidade: -0,0953;
- Endividamento Total: 1,1053
- Rácio de Cobertura de Juros: 8,6548;
- Índice de concentração industrial (HHI): 3

Substituindo os coeficientes na equação (8), obteremos o *Score*:

$$\begin{aligned} \text{Score} = & -9,9469 + 0,3681 * 4,9997 - 0,1424 * -0,0953 + 0,2228 * 1,1053 \\ & - 0,3112 * 8,6548 + 1,4097 * 3 = -6,3110 \quad (10) \end{aligned}$$

Portanto, a probabilidade do *score*, ou seja, da empresa “ONDAH” falir é a seguinte:

$$\text{Prob}(\text{Falência} = 1) = \frac{1}{1 + e^{-(-6.3110)}} = 0.0018 \quad (11)$$

Conforme analisado no subcapítulo 5.4.2, o modelo final de previsão de falência apresenta uma zona de *cutoff* de 1.84%. Como o *score* da empresa que serviu de exemplo é -6.3111 e o mesmo é inferior a 1,84%, podemos concluir que a mesma é classificada como sendo uma empresa saudável.