



ASSOCIAÇÃO DE POLITÉCNICOS DO NORTE (APNOR)
INSTITUTO POLITÉCNICO DE BRAGANÇA

**Combinação da Lei de Benford e de um Modelo Baseado nos
Accruals Discricionários para Aferir a Qualidade da Informação
Financeira**

Érica Braz Barros

Dissertação apresentada ao *Instituto Politécnico de Bragança*
Para obtenção do grau de mestre em Contabilidade e Finanças

Orientação:

Prof. Doutor Jorge Manuel Afonso Alves

Prof. Doutora Ana Isabel Rodrigues Fernandes

Bragança, outubro de 2025



ASSOCIAÇÃO DE POLITÉCNICOS DO NORTE (APNOR)

INSTITUTO POLITÉCNICO DE BRAGANÇA

**Combinação da Lei de Benford e de um Modelo Baseado nos
Accruals Discricionários para Aferir a Qualidade da Informação
Financeira**

Érica Braz Barros

Orientação:

Prof. Doutor Jorge Manuel Afonso Alves

Prof. Doutora Ana Isabel Rodrigues Fernandes

Bragança, outubro de 2025

Resumo

A presente dissertação tem como objetivo avaliar a eficácia da combinação de duas *proxies* utilizadas na literatura para aferir a Qualidade da Informação Financeira (QIF): Lei de Benford e um modelo baseado nos *accruals* discricionários. A combinação destas duas formas de aferir a QIF pode fornecer uma metodologia mais robusta e integrada para aferir a QIF e identificar potenciais manipulações nos relatórios financeiros das empresas. A Lei de Benford é uma técnica estatística utilizada para detetar anomalias numéricas em grandes conjuntos de dados, enquanto que o modelo de *accruals* discricionários, nomeadamente o Modelo de Jones, utiliza os *accruals* para identificar manipulações de resultados. A investigação classificou, numa primeira fase, as empresas do setor da restauração e similares com indícios de manipulação dos resultados, por um lado, com recurso à aplicação da Lei de Benford e, por outro lado, com recurso à aplicação do modelo de Jones. Numa segunda fase, pretendeu-se identificar as empresas com indícios de manipulação dos resultados que sejam classificadas em simultâneo por ambos os métodos. Em cada uma das fases relacionou-se a classificação das empresas com indícios de manipulação de resultados com o incentivo de minimização do pagamento de impostos ao Estado. A metodologia inclui uma análise detalhada dos resultados obtidos pela aplicação de cada técnica individualmente e em conjunto. Em suma, os resultados indicam que a combinação dos dois métodos melhora de forma substancial a capacidade de deteção de manipulação e evidencia indícios localizados de minimização fiscal, demonstrando o potencial destas metodologias como instrumentos complementares de auditoria e controlo fiscal.

Palavras - Chave: Lei de Benford, Modelo de *accruals* discricionários, anomalias, manipulações de resultados

Abstract

The present dissertation aims to evaluate the effectiveness of combining two proxies used in the literature to measure Financial Information Quality (QIF): Benford's Law and a model based on discretionary accruals. The combination of these two ways of measuring QIF provided a more robust and integrated methodology for measuring QIF and identifying potential manipulations in companies' financial reports. Benford's Law is a statistical technique used to detect numerical anomalies in large data sets, while the discretionary accruals model, namely the Jones Model, uses accruals to identify earnings manipulation. The research initially classifies companies in the restaurant and related sectors with signs of earnings manipulation, through the application of Benford's Law on the one hand, and through the application of the Jones model on the other. In a second phase, it was intended to identify companies with signs of earnings manipulation that are classified simultaneously by both methods. In each of the phases, the classification of companies with evidence of earnings manipulation was related to the incentive to minimise tax payments to the State. The methodology includes a detailed analysis of the results obtained by applying each technique individually and in combination. In conclusion, the results indicate that combining the two methods substantially enhances the ability to detect manipulation and provides localised evidence of tax minimisation, demonstrating the potential of these methodologies as complementary instruments for auditing and tax control.

Keywords: Benford's Law, Discretionary accruals model, anomalies, earnings manipulation

Resumen

El objetivo de esta tesis es evaluar la eficacia de la combinación de dos proxies utilizadas en la literatura para medir la calidad de la información financiera (QIF): la ley de Benford y un modelo basado en los devengos discrecionales. La combinación de estas dos formas de medir la QIF proporcionó una metodología más sólida e integrada para medir la QIF e identificar posibles manipulaciones en los informes financieros de las empresas. La Ley de Benford es una técnica estadística utilizada para detectar anomalías numéricas en grandes conjuntos de datos, mientras que el modelo de devengos discrecionales, concretamente el Modelo de Jones, utiliza los devengos para identificar manipulaciones de resultados. La investigación clasifica, en una primera fase, las empresas del sector de la restauración y similares con indicios de manipulación de resultados, por un lado, mediante la aplicación de la Ley de Benford y, por otro, mediante la aplicación del modelo de Jones. En una segunda fase, pretendió identificar las empresas con indicios de manipulación de resultados que sean clasificadas simultáneamente por ambos métodos. En cada una de las fases se se relacionó la clasificación de las empresas con indicios de manipulación de resultados con el incentivo de minimizar el pago de impuestos al Estado. La metodología incluye un análisis detallado de los resultados obtenidos mediante la aplicación de cada técnica individualmente y en conjunto. En resumen, los resultados indican que la combinación de ambos métodos mejora sustancialmente la capacidad de detección de manipulaciones y pone de manifiesto indicios localizados de minimización fiscal, lo que demuestra el potencial de estas metodologías como instrumentos complementarios de auditoría y control fiscal.

Palabras clave: Ley de Benford, modelo de devengos discrecionales, anomalías, manipulaciones de resultados.

Agradecimentos

Aos Professores e Orientadores Jorge Alves e Ana Fernandes, expresso o meu mais profundo agradecimento pela inspiração constante, pela disponibilidade incansável e pela orientação rigorosa e generosa. As vossas sugestões, o vosso espírito crítico e a confiança que depositaram em mim foram determinantes para definir o caminho a seguir e para alcançar os objetivos a que me propus com a escolha deste tema. Cada reunião, comentário e desafio lançado contribuiu para elevar a qualidade deste trabalho e para o meu crescimento académico e pessoal.

À minha família e ao meu namorado, expresso profundo reconhecimento. Aos meus pais, pela educação, confiança e apoio que sustentaram cada etapa; à minha irmã, pela disponibilidade e encorajamento permanentes; e ao meu namorado, pela paciência, serenidade e apoio constante, essenciais para a conclusão desta tese. A todos, o meu sincero obrigada!

Abreviaturas e/ou Acrónimos

A – Total do Ativo

AC – Variação do Ativo Corrente

AFT – Ativos Fixos Tangíveis

AFTD – Efeito Modesto dos Ativos Fixos Tangíveis

AT - *Accruals* Totais

CAE – Código de Atividade Económica

CX – Caixa ou Equivalentes a Caixa

DAC – *Accruals* Discricionários

DACA – *Accruals* Discricionários Absolutos

Dep – Despesa com Depreciação e Amortização

DF - Demonstrações Financeiras

DFC - Demonstração dos Fluxos de Caixa

Dív – Dívida a Curto Prazo

FOC – Financiamentos Obtidos Correntes

GRDA – Gastos e Reversões de Depreciação e de Amortização

IFRS – *International Financial Reporting Standards*

ISRI – Imposto sobre o Rendimento do Período

LBC – Lei de Benford Crítica

NDIM – Dimensão Empresarial

OR – *Odds Ratio*

P – Probabilidade

PC – Passivo Corrente

PIB - Produto Interno Bruto

PME - Pequenas e Médias Empresas

POC - Plano Oficial de Contas

PPE – Propriedade, Planta e Equipamento

QIF – Qualidade da Informação Financeira

RAI – Resultado Antes de Imposto

SABI - Sistema de Análise de Balanços Ibéricos

SARIMA - *Seasonal Autoregressive Integrated Moving Average*

SNC - Sistema de Normalização Contabilístico

TPC – Total do Passivo Corrente

UE - União Europeia

VPS – Vendas e Prestações de Serviços

VEPSD – Variação das Receitas

X – Vetor de Preditores

Índice

Lista de Tabelas	viii
Introdução	9
1. Revisão de Literatura	11
1.1. Qualidade da Informação Financeira.....	11
1.2. Manipulação dos Resultados.....	13
1.3. Lei de Benford	17
1.4. Modelo de <i>Accruals</i> Discricionários (Modelo de Jones)	21
1.5. Combinação de Diferentes Métodos para a Avaliação da Qualidade da Informação Financeira 23	
2. Metodologia de Investigação.....	26
2.1. Objetivo do Estudo e Hipóteses de Investigação	26
2.2. Descrição do Instrumento de Recolha de Dados	27
2.3. População vs. Amostra.....	27
2.4. Descrição dos Métodos de Tratamento dos Dados	28
2.4.1. Aplicação da Lei de Benford.....	28
2.4.2. Estimativa dos <i>Accruals</i> Discricionários	30
3. Apresentação e Análise dos Resultados Empíricos	32
3.1. Caracterização da Amostra	32
3.2. Análise dos Resultados da Aplicação da Lei de Benford	35
3.3. Análise dos Resultados do Modelo Baseado nos <i>Accruals</i>	38
3.4. Análise dos Índicios de Manipulação de Resultados com Incentivo da Minimização do Pagamento de Impostos ao Estado	42
3.5. Comparação e Integração dos Resultados.....	47
3.6. Discussão dos Resultados	52
Conclusões, Limitações e Linhas de Investigação Futuras.....	55
Referências	56
Apêndice	60

Lista de Tabelas

Tabela 1 - Probabilidade de ocorrência para o primeiro dígito	18
Tabela 2 - Probabilidade de ocorrência para o segundo dígito	19
Tabela 3 - Síntese de estudos elaborados no âmbito da aplicação da Lei de Benford	19
Tabela 4 - Evolução da amostra selecionada.....	27
Tabela 5 - Tabela de estatística descritiva	32
Tabela 6 - Frequências do primeiro dígito.....	36
Tabela 7 - Frequências do primeiro e segundo dígito que tomam valores críticos.....	37
Tabela 8 - Regressão linear às variáveis em análise, por ano	40
Tabela 9 - Correlação de Pearson e de Correlação de Spearman por ano	42
Tabela 10 - Correlação de Pearson por ano e por quartil	44
Tabela 11 - Modelo <i>Logit</i> por ano.....	49
Tabela 12 - Modelo <i>Logit</i> por ano e por quartil.....	50

Introdução

A Qualidade da Informação Financeira (QIF) constitui um elemento essencial para a transparência e credibilidade do sistema económico, sustentando a confiança dos investidores, credores, reguladores e autoridades fiscais. A fiabilidade do relato financeiro é determinante para a eficiência dos mercados, para a correta avaliação do desempenho empresarial e para a justiça fiscal. No entanto, a crescente complexidade dos normativos contabilísticos e os incentivos económicos associados ao resultado líquido criam oportunidades e motivações para práticas de manipulação contabilística, afetando a comparabilidade e a utilidade da informação divulgada. Este fenómeno adquire especial relevância em setores marcados por margens reduzidas, forte presença de transações em numerário e elevada atomização empresarial, como é o caso do setor da restauração e similares, onde a tentação de ajustar resultados com vista à minimização da carga fiscal pode ser mais intensa (Christensen et al., 2022; Martins, 2023).

Neste contexto, a literatura tem desenvolvido diferentes metodologias para avaliar a QIF e detetar potenciais indícios de manipulação de resultados. Entre as abordagens mais recorrentes, destacam-se a Lei de Benford, uma ferramenta estatística amplamente utilizada na auditoria e na deteção de fraudes para identificar anomalias na distribuição dos dígitos de grandes conjuntos de dados (Swanson et al., 2003; Silva et al., 2017), e os modelos baseados nos *accruals* discricionários, que procuram quantificar a componente discricionária dos ajustamentos contabilísticos, sendo o Modelo de Jones (1991) uma das suas formulações mais reconhecidas. A aplicação da Lei de Benford permite captar padrões numéricos inesperados em registos contabilísticos, enquanto os modelos de *accruals* avaliam o grau de discricionariedade na gestão dos resultados, oferecendo perspetivas complementares sobre a qualidade da informação.

Apesar da sua ampla utilização individual, a aplicação conjunta destas duas metodologias permanece escassa. A literatura raramente aborda de forma integrada a combinação entre uma técnica estatística de deteção de anomalias e um modelo contabilístico de estimação de *accruals* discricionários, deixando em aberto a questão de saber se a convergência entre ambas pode aumentar a robustez da avaliação da QIF e melhorar a capacidade de identificação de indícios de manipulação de resultados. É precisamente essa lacuna que o presente estudo procura colmatar, propondo-se a avaliar a eficácia da combinação da Lei de Benford com o Modelo de Jones (1991) na aferição da QIF e na identificação de potenciais manipulações em relatórios de empresas do setor da restauração e similares.

O objetivo principal desta investigação consiste em avaliar se a combinação destas duas *proxies* de QIF permite uma melhor classificação das empresas com indícios de manipulação de resultados, particularmente quando esta manipulação possa estar associada a incentivos de minimização da carga fiscal. Parte-se da premissa de que sinais convergentes provenientes de métodos

conceptualmente distintos, um baseado em padrões numéricos e outro em relações contabilísticas, poderão representar um indício mais fiável de práticas de manipulação de resultados do que a aplicação isolada de cada método.

A metodologia do estudo divide-se em duas fases principais. Na primeira, aplicam-se separadamente as duas técnicas: a Lei de Benford, para detetar irregularidades nos padrões numéricos de rubricas como vendas e contas a receber, e o Modelo de Jones (1991), para estimar os *accruals* discricionários e identificar possíveis anomalias contabilísticas. Na segunda fase, os resultados obtidos com ambos os métodos são cruzados, sendo consideradas com indícios de manipulação apenas as empresas que são sinalizadas simultaneamente pelos dois. Esta abordagem conjunta permite aumentar a fiabilidade da deteção e reduzir o risco de classificações incorretas, proporcionando uma análise mais robusta da QIF.

Para complementar a análise, o estudo relaciona os indícios de manipulação identificados com medidas de incentivo fiscal, nomeadamente indicadores associados à minimização do imposto sobre o rendimento. Testam-se associações estatísticas entre as classificações obtidas (Lei de Benford, Modelo de Jones e combinação de ambos) e variáveis como o imposto sobre o resultado do período (ISR), procurando aferir se as empresas que apresentam padrões anómalos também exibem comportamentos consistentes com gestão fiscal agressiva. Desta forma, a investigação procura compreender não apenas a eficácia metodológica da combinação proposta, mas também as suas implicações económicas e fiscais.

Os contributos esperados deste estudo são de natureza metodológica, empírica e prática. No plano metodológico, propõe-se um enquadramento integrado que combina duas técnicas consolidadas de aferição da QIF, proporcionando uma ferramenta mais robusta para a identificação de anomalias contabilísticas. No plano empírico, o trabalho acrescenta evidência recente sobre o setor da restauração e similares, num contexto particularmente relevante e ainda pouco explorado na literatura. No plano prático, os resultados poderão apoiar auditores, reguladores e autoridades fiscais no desenvolvimento de procedimentos de triagem e de auditoria analítica mais eficientes, orientados para a deteção de manipulação e evasão fiscal.

A estrutura da dissertação reflete este percurso analítico. Após esta introdução, a secção seguinte apresenta a revisão de literatura sobre a QIF, manipulação de resultados, Lei de Benford e Modelo de Jones, culminando na proposta da combinação dos dois métodos. A metodologia de investigação, incluindo os objetivos, hipóteses, população, amostra e métodos de tratamento de dados. Após a secção dedicada à metodologia, a apresentam-se e analisam-se os resultados empíricos, comparando as classificações geradas por cada método e avaliando a sua relação com incentivos fiscais. A penúltima secção discute os resultados à luz da literatura existente, enquanto a última secção sintetiza as conclusões, limitações e linhas de investigação futura.

1. Revisão de Literatura

A QIF é um fator determinante para a transparência e fiabilidade dos relatórios contabilísticos, desempenhando um papel essencial na tomada de decisões por parte de investidores, reguladores e demais partes interessadas (Achim & Chis, 2014; Renkas et al., 2016). No entanto, a integridade desta informação pode ser comprimida pela manipulação dos resultados, um fenómeno que ocorre quando os gestores ajustam deliberadamente os dados financeiros para apresentar uma imagem mais favorável da organização. Esta prática pode distorcer a perceção do desempenho financeiro, influenciando decisões económicas de forma indevida (Christensen et al., 2022; Elisabeta et al., 2014; Trussel, 2003).

Apesar da existência de mais métodos, o foco na deteção de manipulação contabilística vai apoiar-se sobretudo na Lei de Benford e no Modelo de Jones (Antunes et al., 2023; Jones, 1991). A combinação destes métodos reforça a credibilidade da análise, cruzando as anomalias numéricas com padrões de decisão da administração (Licerán-Gutiérrez & Cano-Rodríguez, 2019; Trussel, 2003).

1.1. Qualidade da Informação Financeira

A QIF é um conceito multifacetado que abrange diversas características essenciais para a tomada de decisões. Os principais atributos incluem relevância, fiabilidade, comparabilidade e compreensibilidade (Comissão de Normalização Contabilística [CNC], 2021). Estas características qualitativas são delineadas em estruturas conceptuais pelos principais organismos de normalização contabilística, embora as definições possam variar entre países (Achim & Chiş, 2014; Renkas et al., 2016).

Além disso, a QIF desenvolve-se em três níveis: representação da realidade, atribuição de significado e transmissão da informação (Dinu, 2022). Para melhorar a utilidade dos relatórios financeiros, podem ser aplicadas características amplificadoras, desde que os dados sejam relevantes e fidedignos. A comparabilidade é um desses elementos essenciais, permitindo identificar semelhanças e diferenças entre Demonstrações Financeiras (DF) de várias entidades ou ao longo do tempo dentro da mesma organização. Para assegurar essa comparabilidade, é fundamental manter a consistência na aplicação das políticas contabilísticas (Dinu, 2022).

Embora ainda não exista uma definição universalmente aceite, a QIF é fundamental para a manutenção da confiança das partes interessadas e para dar suporte a análises financeiras complexas (Achim & Chiş, 2014).

A QIF é um conceito que, apesar da sua importância no âmbito económico e empresarial, não possui uma definição universal e concreta. Este conceito tem sido explorado a partir de diferentes

perspetivas, refletindo a sua natureza multifacetada e a sua relevância para a transparência e a tomada de decisão no mercado. No domínio filosófico, a qualidade refere-se às propriedades objetivas dos fenómenos do mundo real, percebidas através dos sentidos. No entanto, no contexto das ciências económicas, o conceito de qualidade inicialmente estava associado a bens materiais, mas com o tempo passou a abranger também a informação financeira (Renkas et al., 2016).

A contabilidade desempenha um papel crucial na produção da informação financeira, transformando os dados em produtos informacionais com valor económico. Com a evolução das *International Financial Reporting Standards* (IFRS), o conceito de QIF foi formalmente reconhecido como uma categoria económica, estabelecendo critérios para garantir a fiabilidade, a relevância e a comparabilidade dos relatórios financeiros (Jara et al., 2011; Renkas et al., 2016). Estes critérios são fundamentais para assegurar que a informação divulgada pelas empresas seja útil para a tomada de decisões por parte de investidores, reguladores e gestores. A implementação das IFRS, por sua vez, tem sido apontada como um fator determinante na melhoria da QIF, ao promover a uniformidade e a transparência nos relatórios corporativos (Jara, et al., 2011).

A literatura destaca a importância da QIF na avaliação do desempenho empresarial e na construção da confiança no mercado. Uma informação financeira de qualidade é crucial para garantir a transparência e facilitar a tomada de decisões fundamentadas pelos diversos *stakeholders* (Christensen et al., 2022). Além disso, fatores como a responsabilidade social corporativa têm-se tornado cada vez mais relevantes na determinação da QIF, fazendo parte atualmente, dos planos estratégicos corporativos, desenvolvidos pelos gestores. Um compromisso mais profundo com a responsabilidade social tem demonstrado ter um impacto positivo na QIF, uma vez que incentiva práticas mais transparentes e éticas, gerando benefícios a longo prazo para os *stakeholders* (Christensen et al., 2022; Yan, 2024).

Estudos empíricos também sugerem que a QIF pode variar significativamente consoante o setor de atividade e o ambiente económico. Por um lado, em determinadas regiões, como na Tunísia, verificou-se que a QIF em algumas organizações é reduzida, especialmente em setores não financeiros. Por outro lado, empresas do setor financeiro apresentam níveis mais elevados de provisões discricionárias, mas apresentam uma menor probabilidade de reformular as DF em comparação com outros setores (Aloui & Belanès, 2024). Estas diferenças evidenciam como fatores institucionais e setoriais podem influenciar a qualidade dos relatórios financeiros.

A QIF pode ser influenciada por diversos fatores internos e externos às organizações. Elementos como a administração organizacional, o setor de atividade e as características qualitativas dos relatórios de gestão desempenham um papel fundamental na precisão e fiabilidade da informação financeira. Além disso, a literacia financeira, o nível de escolaridade, a utilização de tecnologias da informação e a adoção de normas contabilísticas específicas têm um impacto direto na qualidade dos

relatórios, reforçando a importância da capacitação e da modernização dos processos (Budianto et al., 2024).

Outro fator relevante para a QIF é a adoção de tecnologias de informação e de normas contábilísticas específicas. A implementação de ferramentas tecnológicas avançadas e a conformidade com padrões contábilísticos rigorosos demonstraram melhorar substancialmente a precisão e a fiabilidade dos relatórios financeiros. Esta relação é particularmente evidente em Pequenas e Médias Empresas (PME) do setor da restauração, onde a utilização adequada da tecnologia e a aplicação das IFRS resultam em ganhos significativos na QIF (Budianto et al., 2024).

Deste modo, a QIF é influenciada por um conjunto diversificado de fatores, incluindo as normas contábilísticas, a responsabilidade social corporativa, o setor de atividade e a tecnologia da informação. Estes elementos, em conjunto, determinam a fiabilidade, a relevância e a utilidade da informação financeira, sendo fundamentais para assegurar a transparência e a confiança nos mercados financeiros. A sua importância transcende a mera conformidade legal, desempenhando um papel estratégico na gestão empresarial e na criação de valor a longo prazo.

1.2. Manipulação dos Resultados

A manipulação de resultados consiste na alteração deliberada das informações financeiras de uma organização, com o objetivo de apresentar uma imagem financeira mais favorável do que a realidade. Este fenómeno ocorre tanto no setor lucrativo como no setor não lucrativo e representa uma preocupação significativa para investidores, reguladores e demais partes interessadas (Elisabeta et al., 2014; Trussel, 2003). Em essência, a manipulação de resultados traduz-se em práticas contábilísticas que, embora possam estar em conformidade com as normas, distorcem a verdadeira situação financeira da entidade (Christensen et al., 2022).

As principais motivações para a manipulação dos resultados são variadas e estão geralmente relacionadas com objetivos estratégicos e interesses específicos dos gestores. Entre estas motivações destacam-se: influenciar a perceção dos investidores para aumentar o valor de mercado da empresa, maximizar a remuneração variável dos gestores baseada em resultados financeiros, melhorar as condições para a obtenção de crédito e reduzir o risco de supervisão e sanções por parte das entidades reguladoras (Cupertino & Martinez, 2008). Além disso, a manipulação de resultados pode também estar associada à tentativa de cumprir metas de desempenho estabelecidas ou de evitar a divulgação de resultados desfavoráveis que possam prejudicar a reputação da organização (Elisabeta et al., 2014).

Trussel (2003), destaca a necessidade de ferramentas para avaliar as informações financeiras e propõe indicadores como margem de excedente, taxa de despesas diferidas, crescimento de receita, taxa de depreciações e a mudança na proporção de gastos com programas.

No setor não lucrativo, a manipulação de resultados pode ocorrer com o intuito de demonstrar maior eficiência na aplicação de recursos. Por exemplo, algumas organizações podem inflacionar a proporção de despesas em programas essenciais como assistência social, educação, saúde, meio ambiente e direitos humanos, em relação ao total das despesas, criando uma percepção de maior compromisso com a sua missão institucional (Trussel, 2003). Este tipo de manipulação procura reforçar a confiança de doadores e financiadores, assegurando a continuidade do financiamento e a legitimidade institucional.

No contexto da auditoria, a identificação de sinais de risco, conhecidos como "*red flags*", é um procedimento fundamental na fase de pré-planeamento. Este processo envolve a análise de indícios que possam sugerir práticas de manipulação, como variações anômalas em indicadores financeiros, incongruências nas DF e a presença de incentivos para manipular resultados (Cupertino & Martinez, 2008). Adicionalmente, os procedimentos de revisão analítica permitem a comparação de valores absolutos, a análise de índices financeiros e a avaliação de tendências temporais, facilitando a identificação de anomalias (Arens et al., 2017).

Fatores externos também desempenham um papel relevante na prática de manipulação de resultados. Valaskova e Fedorko (2021) destacam a influência de variáveis como a corrupção, a taxa de impostos, a burocracia e as regulamentações tributárias. Embora não seja possível determinar com precisão o impacto isolado de cada fator no número de empresas que recorrem à manipulação, é evidente que sistemas fiscais complexos e ambientes regulatórios pouco transparentes favorecem tais práticas (Valaskova & Fedorko, 2021).

No que diz respeito aos regimes fiscais, Pais e Dias (2022) exploram a relação entre o regime simplificado e a conformidade entre o lucro contábilístico e o lucro tributável, analisando os incentivos para gerir os resultados em diferentes regimes fiscais. Para além desta correlação, também exploram o impacto do regime simplificado na conformidade tributária e nas práticas de manipulação de resultados, evidenciando como as alterações nos regimes fiscais podem influenciar estas práticas (Pais & Dias, 2022). Os autores concluem que as empresas no regime simplificado tendem a manipular menos os resultados em comparação com aquelas sujeitas ao regime geral. No entanto, observam que o regime simplificado incentiva o aumento dos lucros, enquanto o regime geral tende a reduzi-los. Ademais, constatam que os ajustes com elevada conformidade fiscal são frequentemente utilizados para reduzir o rendimento tributável, ao passo que ajustes de baixa conformidade servem para aumentar o lucro contábilístico (Pais & Dias, 2022). Por conseguinte, uma menor conformidade fiscal está associada a uma redução na manipulação de lucros (Pais & Dias, 2022).

Dado que a manipulação de resultados compromete a transparência e enfraquece a confiança dos investidores, torna-se imperativo o desenvolvimento contínuo de métodos robustos para detetar e mitigar estas práticas (Licerán-Gutiérrez & Cano-Rodríguez, 2019). A utilização de modelos

quantitativos, aliada a uma abordagem analítica rigorosa e a uma auditoria independente e eficaz, desempenha um papel crucial na prevenção e na identificação de manipulações financeiras, assegurando, assim, maior integridade e fiabilidade na divulgação das informações financeiras.

Considerando que nenhum método isolado consegue abranger na íntegra os diversos aspetos da QIF, adota-se cada vez mais abordagens mistas, que combinam técnicas qualitativas e quantitativas de acordo com o enquadramento dos dados disponíveis, permitindo assim avaliar de modo integral a fiabilidade, a clareza e a relevância das DF (Licerán-Gutiérrez & Cano-Rodríguez, 2019). De entre os métodos mais comuns, destacam-se (Cupertino & Martinez, 2008):

- Os **métodos estatísticos**, que utilizam análises estatísticas para identificar anomalias nos dados, como a Lei de Benford, os testes Qui-Quadrado e a análise de regressão. Estes métodos são eficazes na deteção de padrões atípicos que possam indicar fraudes ou manipulação.
- A **revisão analítica** envolve a avaliação das informações financeiras por meio de relações plausíveis entre dados financeiros e não financeiros, permitindo identificar sinais de alerta, ou *red flags*, que possam indicar comportamentos suspeitos.
- Já os **modelos de aprendizagem automática**, como os algoritmos de classificação e *clustering*, são utilizados para analisar grandes volumes de dados e detetar padrões ou anomalias que possam ser indicativos de fraude.
- A **análise de accruals** foca na investigação dos *accruals* discricionários, com o objetivo de identificar manipulação de resultados contabilísticos, utilizando modelos como o *Beneish Model*, que é amplamente utilizado para este fim. Além disso, o Modelo de Jones, também é uma ferramenta essencial na análise de *accruals*, pois permite separar os *accruals* normais dos discricionários, ajudando a identificar anomalias que podem ser indicativas de manipulação de resultados financeiros.
- A **auditoria forense** combina técnicas de auditoria tradicionais com investigações detalhadas, com o intuito de identificar fraudes em demonstrações financeiras, muitas vezes em situações mais complexas.
- A **análise de redes** examina as relações entre diferentes entidades, como transações entre empresas, para identificar comportamentos suspeitos ou padrões que possam indicar fraude. As técnicas de *data mining* utilizam algoritmos para explorar grandes conjuntos de dados em busca de padrões que possam sugerir atividades fraudulentas.

Essas metodologias podem ser usadas de forma isolada ou combinada, conforme o caso, com o objetivo de aumentar a eficácia na deteção de fraudes e garantir a integridade das informações financeiras (Cupertino & Martinez, 2008).

Para detecção de manipulação de resultados em conjuntos de dados, recorre-se atualmente a técnicas estatísticas de distribuição e de séries temporais: a Lei de Benford analisa o padrão esperado dos dígitos iniciais em dados numéricos, sobretudo em relatórios financeiros, sinalizando desvios anómalos na frequência dos dígitos significativos, enquanto a análise de séries temporais utiliza modelos autorregressivos e *Seasonal Autoregressive Integrated Moving Average* (SARIMA) - modelo estatístico utilizado na análise e previsão de séries temporais - para prever comportamentos sazonais e modelos lineares generalizados para dados de contagem, permitindo identificar flutuações que se afastam do padrão esperado ao longo do tempo (Antunes et al., 2023; Latorre & Cardoso, 2001; Oliveira et al., 2014).

Paralelamente, aplicam-se modelos de regressão que estimam ajustes contabilísticos intencionais, como o Modelo de Jones, que calcula *accruals* discricionários e evidencia discrepâncias entre os resultados reportados e os previstos, sinalizando potenciais fraudes contabilísticas, e técnicas de detecção de *outliers* (*Z-score*, teste de *Grubbs*, *boxplots*) para destacar valores isolados que se afastam do comportamento estatístico global (Jones, 1991; Oliveira et al., 2014; Silva & Nóbrega, 2021).

Para reforçar a fiabilidade do diagnóstico, combinam-se ainda abordagens de *machine learning*, que identificam padrões anómalos em grandes volumes de dados, análise de redes para expor transações suspeitas e perícia forense digital para investigar alterações nos sistemas de registo (Antunes et al., 2023; Mendonça et al., 2022).

Para além destas medidas para identificar a manipulação dos resultados, Licerán-Gutiérrez e Cano-Rodríguez (2019) abrangem um conjunto de *proxies* empíricas para identificar manipulação de resultados em DF, distribuídas em quatro dimensões principais.

Primeiramente, a gestão discricionária de resultados (*earnings management*) recorre a *accruals* anormais, que são obtidos pela diferença entre *accruals* reportados e previstos por modelos como Jones ou Dechow, bem como à análise da concentração de lucros ao redor de metas, sinalizando ajustes contabilísticos intencionais.

Em segundo lugar, o alisamento de resultados (*earnings smoothing*) é mensurado pela razão entre volatilidade do lucro e volatilidade de vendas ou de fluxo de caixa operacional e pela correlação entre *accruals* discricionários e variações de caixa, uma vez existentes correlações excessivamente elevadas sugerem compensações artificiais para estabilizar o lucro.

A terceira dimensão explora propriedades de séries temporais dos lucros, sobretudo a sua persistência (via coeficientes autorregressivos) e variância, onde a baixa previsibilidade ou volatilidade anómala indicam ruído introduzido por práticas manipulativas.

Por fim, os indicadores de conservadorismo contábilístico são diferenciados entre conservadorismo condicional, medido por modelos como o de Basu ou o *C-Score de Khan e Watts*, que avaliam a rapidez no reconhecimento de perdas face a ganhos, e conservadorismo incondicional medido pelo *market-to-book* ou saldos acumulados de *accruals* negativos, o que evidencia a existência de reservas ocultas e práticas de avaliação deliberadamente mais prudentes. A utilização conjunta destas diversas *proxies* permite obter um retrato mais sólido da qualidade da informação contábilística e identificar eventuais desvios intencionais nas práticas de relato financeiro (Licerán-Gutiérrez & Cano-Rodríguez, 2019).

De acordo com Trussel (2003), são necessários diversos testes para se obter uma conclusão mais robusta, contudo, neste caso, a combinação da Lei de Benford com o Modelo de Jones que permite um enquadramento mais sólido na identificação de manipulações financeiras. Ao integrar estas metodologias, torna-se possível detetar tanto irregularidades numéricas nos registos contábilísticos como tendências atípicas na gestão dos resultados, fornecendo um enquadramento robusto para a monitorização da QIF e o reforço da confiança nos relatórios das organizações (Licerán-Gutiérrez & Cano-Rodríguez, 2019).

A Lei de Benford baseia-se na análise da frequência dos dígitos significativos em conjuntos de dados financeiros, assumindo que, em circunstâncias normais, os números seguem uma distribuição estatística previsível. Os desvios significativos em relação a essa distribuição podem indicar a existência de manipulação (Licerán-Gutiérrez & Cano-Rodríguez, 2019). Já o Modelo de Jones (1991) propõe a utilização de *accruals* discricionários para aferir a manipulação, distinguindo entre os *accruals* normais e os gerados por práticas intencionais de gestão de resultados.

Desta forma, irá ser aprofundado de seguida a Lei de Benford e o Modelo de Jones, evidenciando a sua integração na monitorização contínua da QIF.

1.3. Lei de Benford

A fraude é um dilema cada vez mais comum no dia a dia que tem afetado a sociedade, desde indivíduos a organizações, sendo estas práticas consideradas um crime (Antunes et al., 2023). Com o objetivo de prevenir e mitigar esta manipulação de resultados ou fraude, a Lei de Benford pode ser aplicada como uma ferramenta estatística para identificar padrões anómalos nos dados, permitindo uma análise mais rigorosa e a deteção de discrepâncias que possam indicar práticas fraudulentas (Antunes et al., 2023).

De acordo com Silva et al. (2017), a Lei de Benford foi inicialmente desenvolvida por Simon Newcomb, em 1881, e redescoberta por Frank Benford, relativamente à qual ambos concluíram que a

probabilidade da ocorrência do primeiro dígito é mais elevada do que a ocorrência do último dígito. Desta forma, esta lei tem sido utilizada para identificar anomalias numéricas, ou seja, utilizam-na em auditoria, para verificar se há desvios de dados financeiros nas empresas (Antunes et al., 2023). Porque, uma vez manipulados os dados, estes deixam de seguir a probabilidade esperada pela Lei de Benford (Antunes et al., 2023; Silva et al., 2017). A Lei de Benford surgiu a partir da constatação de que as primeiras páginas das tabelas logarítmicas, que continham números iniciados pelo dígito 1, estavam mais desgastadas do que as últimas (Crispim, 2023; Silva et al., 2017). Frank Benford referia que nos números, os dígitos de 1 e 2, eram mais frequentes do que os dígitos 8 e 9. Desta forma, este estudou a equação [1], ou seja, a probabilidade de um dígito surgir como um primeiro dígito é determinada pela seguinte equação (Benford, 1938; Crispim, 2023; Hill, 1995; Silva et al., 2017):

Probabilidade de ocorrência do primeiro dígito:

$$P(d_1) = \log_{10} \left(1 + \frac{1}{d_1} \right) \quad [1]$$

Nota: P(d) - probabilidade; d₁ - primeiro dígito de um número.

A Tabela 1 apresenta a probabilidade de ocorrência para o primeiro dígito.

Tabela 1 - Probabilidade de ocorrência para o primeiro dígito

d ₁	1	2	3	4	5	6	7	8	9
P(d ₁)	0,30103	0,17609	0,12494	0,09691	0,07918	0,06695	0,05799	0,05115	0,04576

Fonte: Elaboração própria.

Como se pode observar na Tabela 1, a probabilidade de sair um primeiro dígito com um valor mais baixo é maior que a probabilidade de sair um primeiro dígito com um valor mais alto.

De seguida, observa-se a probabilidade do segundo dígito, que é determinada pela equação [2] (Antunes et al., 2023; Benford, 1938; Crispim, 2023; Hill, 1995; Silva et al., 2017):

Probabilidade de ocorrência dos dois primeiros dígitos:

$$P(d_1 d_2) = \log \left(1 + \frac{1}{d_1 d_2} \right) \quad [2]$$

Nota: P(d) - probabilidade; d₁ – primeiro dígito de um número; d₂ - segundo dígito de um número.

A Tabela 2 apresenta a probabilidade de ocorrência para o segundo dígito:

Tabela 2 - Probabilidade de ocorrência para o segundo dígito

d_2	0	1	2	3	4	5	6	7	8	9
$P(d_2)$	0,11968	0,11389	0,10882	0,10433	0,10031	0,09668	0,09337	0,09035	0,08757	0,08500

Fonte: Elaboração própria.

Como se pode observar na Tabela 2, a probabilidade de sair um segundo dígito com um valor mais baixo é maior que a probabilidade de sair um segundo dígito com um valor mais alto. Logo, tanto na Tabela 1, como na Tabela 2, verifica-se a redescoberta feita por Frank Benford, mas nem sempre um conjunto de dados segue a distribuição esperada, uma vez que há certas restrições associadas à aplicação desta técnica de análise (Antunes et al., 2023).

Vários estudos empíricos têm demonstrado a eficácia da Lei de Benford na auditoria e na contabilidade forense, identificando padrões suspeitos em relatórios financeiros de empresas e instituições. A tabela abaixo apresenta uma síntese de alguns estudos relevantes, destacando os seus objetivos, bases de dados analisadas e principais conclusões.

Tabela 3 - Síntese de estudos elaborados no âmbito da aplicação da Lei de Benford

Autores e ano	Objetivo de estudo	Base de dados	Principais conclusões
Nigrini e Mittermaier (1997)	Explorar como a Lei de Benford pode ser utilizada como ferramenta analítica para detetar irregularidades em dados contabilísticos.	Dados financeiros de empresas.	O estudo mostra que a Análise Digital (DA), baseada na Lei de Benford, ajuda na deteção de fraudes em auditorias. Um caso real identificou padrões anómalos, levando à investigação de um fornecedor suspeito. O DA pode fortalecer a auditoria ao validar a confiabilidade dos dados financeiros.
Lesperance et al., (2016)	Avaliar diferentes testes de aderência à Lei de Benford e a sua aplicabilidade em conjuntos de dados financeiros simulados.	Conjuntos de dados financeiros simulados.	O estudo oferece <i>insights</i> sobre a robustez da técnica da Lei de Benford na deteção de irregularidades em dados financeiros, demonstrando a eficácia dos testes de aderência para identificar padrões não naturais em dados.
Gorenc (2019)	O estudo procura compreender como esta lei pode ser utilizada para identificar anomalias e comportamentos fraudulentos, especialmente no contexto de auditorias e fraudes digitais.	Informações coletadas de diversas fontes, como artigos académicos e estudos de caso	A Lei de Benford é uma ferramenta eficaz e de baixo custo para detetar fraudes, mas deve ser usada com outras técnicas para maior precisão. A deteção precoce é essencial, embora a Lei tenha limitações, como falsos positivos e dependência do tamanho dos dados.

Fonte: Elaboração própria.

A primeira restrição associada à aplicação desta técnica analítica diz respeito ao facto de ser fundamental que os números se distribuam de forma aleatória (Antunes et al., 2023). Para além desta,

uma outra restrição está relacionada com a dimensão e a diversidade do conjunto de dados, uma vez que a Lei de Benford exige que cada registo numérico apresente vários dígitos, a fim de assegurar uma conformidade adequada com a distribuição esperada (Antunes et al., 2023). Esta lei também tem como restrição, poder considerar gerar falsos positivos, na qual dados legítimos são considerados anómalos, conduzindo a investigações desnecessárias (Antunes et al., 2023). Além desta, também a interpretação dos resultados requer experiência e conhecimento, uma vez que esta lei deve ser utilizada apenas como um indicador preliminar e não como uma prova definitiva de fraude (Antunes et al., 2023). A eficácia desta lei, pode variar em função do contexto, dependendo do setor ou do tipo de dados analisados, o que pode exigir ajustes na metodologia (Antunes et al., 2023).

A aplicação desta lei depende do número de dados, mas não há consenso sobre o valor mínimo necessário (Antunes et al., 2023). Enquanto alguns autores sugerem 500 ou 1000 dados, outros referem 300 dados como aceitável, apesar de que a lei já se revelou válida para conjuntos menores (Antunes et al., 2023). É importante salientar que, quanto menor o número de observações, maior a probabilidade de distorção dos dados, sendo ideal um conjunto suficientemente grande para minimizar essa distorção, neste caso, que esta tenda para zero (Antunes et al., 2023; Silva et al., 2017).

Para detetar esta distorção nos dados, utilizam-se o teste Z e o teste do qui-quadrado (χ^2). O teste Z serve para verificar se os dados seguem uma distribuição normal, representada pela seguinte fórmula (Silva et al., 2017):

Equação do teste Z:

$$Z = \frac{PO_i - PE_i}{\sqrt{PE(1 - p_{ei})}} \quad [3]$$

Nota: PO_i – Frequências observadas; p_{ei} – Probabilidade esperada; PE – é o tamanho da amostra (n) a multiplicar pela probabilidade esperada. Fonte: Silva et al. (2017).

Já o teste χ^2 , serve para avaliar a discrepância entre as frequências observadas e as frequências esperadas, representada pela seguinte fórmula (Silva et al., 2017):

Equação do teste do χ^2 :

$$\chi = \frac{(PO_i - PE_i)^2}{PE_i} \quad [4]$$

Nota: PO_i – Frequências observadas; p_{ei} – Probabilidade esperada; PE – é o tamanho da amostra (n) a multiplicar pela probabilidade esperada. Fonte: Elaboração própria, com recurso a Silva, et al., (2017).

A Lei de Benford é uma ferramenta eficaz na deteção de fraudes ou erros, mas não oferece uma prova conclusiva, ou seja, exige investigação complementar. Assim, ao ser combinada com os testes referidos acima, possibilita uma análise mais rigorosa, promovendo transparência e fortalecendo a confiança em auditorias e avaliações estatísticas (Crispim, 2023).

1.4. Modelo de *Accruals* Discricionários (Modelo de Jones)

A manipulação de resultados pode traduzir várias motivações, designadamente influenciar a opinião dos investidores, maximizar a remuneração dos gestores, ou melhorar as condições para obter crédito (Cupertino & Martinez, 2008).

De acordo com Cupertino e Martinez (2008), os *accruals* são um elemento fundamental do modelo contabilístico e podem influenciar a avaliação da eficiência empresarial, retida no resultado líquido.

A informação financeira é frequentemente utilizada para projeções de lucros, ajudando na tomada de decisões de investimento (Cupertino & Martinez, 2008). A diferença entre o resultado contabilístico e o fluxo de caixa deve-se a normas contabilísticas que determinam o reconhecimento de rendimentos e gastos independentemente da transação monetária efetiva, ou seja, alguns rendimentos podem ser registados antes do seu recebimento real, enquanto os gastos, como a depreciação, são contabilizadas sem representarem uma saída de fundos (Cupertino & Martinez, 2008).

A análise da gestão de resultados centra-se nas decisões discricionárias dos gestores que influenciam as DF (Cupertino & Martinez, 2008; Jones, 1991). Para isso, recorre-se a modelos de estimativa que variam em complexidade e pressupostos (Cupertino & Martinez, 2008).

Os modelos mais simples consideram os *accruals* totais (AT), enquanto abordagens mais avançadas como o Modelo de Jones (1991), distinguem a componente discricionária da não discricionária (Cupertino & Martinez, 2008; Jones, 1991). Uma forma clássica de identificar a gestão de resultados, assenta na avaliação dos *accruals*, os quais representam a discrepância entre o fluxo de caixa operacional e o resultado contabilístico (Cupertino & Martinez, 2008).

Os *accruals* dividem-se em não discricionários, determinados por normas contabilísticas, e os discricionários, resultantes das decisões dos gestores (Cupertino & Martinez, 2008). O estudo foca-se nos *accruals* discricionários, onde ocorre a gestão de resultados (Cupertino & Martinez, 2008).

O Modelo de Jones, conforme descrito por Costa e Soares (2022) e por Jones (1991), é uma metodologia baseada na identificação dos *accruals* discricionários, isto é, aqueles ajustes que o gestor faz de modo a manipular os resultados, distinguindo-os dos não discricionários, que ocorrem naturalmente. Para isso, o modelo estima os *accruals* esperados e subtrai-os dos AT, permitindo identificar potenciais manipulações nos resultados financeiros.

De acordo com Jones (1991), os AT são determinados pela diferença entre a variação do fundo de maneiço não monetário, antes da dedução do imposto sobre o rendimento a pagar, e os gastos de depreciação. Esses *accruals* podem ser calculados através do balanço ou da demonstração dos fluxos de caixa (DFC), sendo que, nesta última, resultam da diferença entre o lucro líquido e o fluxo de caixa operacional (Cupertino & Martinez, 2008). O Modelo de Jones, é utilizado para distinguir a componente

discricionária dos AT, ao estabelecer a relação entre os *accruals* não discricionários, que ocorrem naturalmente, e o valor total dos *accruals* (Costa et al., 2018).

Expressão para calcular os AT através do balanço:

$$AT = (\Delta AC - \Delta Cx) - (\Delta PC - \Delta Dív) - Dep \quad [5]$$

Nota: ΔAC - Variação do ativo corrente; ΔCx - Variação do caixa ou equivalentes a caixa; ΔPC - Variação do passivo corrente; $\Delta Dív$ - Variação da dívida a curto prazo; Dep - Despesa com depreciação e amortização. Fonte: Elaboração própria, com recurso a Cupertino e Martinez (2008).

O Modelo de Jones é uma ferramenta utilizada em contabilidade e finanças, para detetar a manipulação de resultados nas DF, através da gestão de resultados (Jones, 1991).

De acordo com Cupertino e Martinez (2008), Jones (1991) e Defond e Subramanyam (1998), o Modelo de Jones expressa os *accruals* em função da variação da receita de vendas e do saldo do ativo fixo tangível. Defond e Subramanyam (1998) ainda referem que estas variáveis controlam os efeitos das mudanças económicas, distinguindo-as dos *accruals* manipulados. A alteração nas receitas tem impacto nos *accruals* relacionados com o capital circulante, enquanto os ativos fixos afetam os *accruals* resultantes da depreciação (Defond & Subramanyam, 1998). Os *accruals* discricionários resultam da diferença entre os AT e as componentes dos *accruals* não discricionários (Defond & Subramanyam, 1998).

Este modelo apresenta como limitações, poder precisar de ajustes para setores específicos com características únicas, e nem todos os acréscimos na manipulação dos resultados resultam de manipulação de resultados intencional (Jones, 1991).

A fórmula do Modelo de Jones para detetar *accruals* discricionários é a seguinte (Costa & Soares, 2022; Jones, 1991):

Fórmula do Modelo de Jones:

$$\left(\frac{TA_{it}}{A_{it-1}}\right) = \alpha_0 + \alpha_1 \left(\frac{1}{A_{it-1}}\right) + \beta_1 \left(\frac{\Delta R_{it}}{A_{it-1}}\right) + \beta_2 \left(\frac{PPE_{it}}{A_{it-1}}\right) + \epsilon_{it} \quad [6]$$

Nota: TA_{it} - Acréscimos totais no ano t para a empresa i; ΔR_{it} - Receitas no ano t menos receitas no ano t-1 para a empresa i; PPE_{it} - Ativos Fixos brutos, instalações e equipamentos no ano t para a empresa i; A_{it-1} - Ativos totais no ano t-1 para a empresa i; ϵ_{it} - Termo de erro no ano t para a empresa i. Fonte: Elaboração própria, com recurso a Costa e Soares, (2022) e Jones (1991).

A variável dependente, acréscimos totais, é calculada pela diferença entre o lucro líquido o fluxo de caixa operacional (Jones, 1991). O resíduo representado pelo erro (ϵ), no seu valor absoluto, reflete os níveis de *accruals* discricionários (Costa et al., 2018).

Neste método, a componente discricionária dos acréscimos é utilizada para identificar a gestão de resultados, uma vez que esta abordagem se justifica pelo facto de essa componente permitir captar uma proporção mais significativa das possíveis manipulações praticadas pelos gestores (Jones, 1991).

Desta forma, é pretendido comparar estes dois métodos e verificar se a sua combinação pode apresentar um potencial significativo, para aprimorar a precisão na deteção de manipulações financeiras.

1.5. Combinação de Diferentes Métodos para a Avaliação da Qualidade da Informação Financeira

A avaliação da QIF constitui um domínio complexo e multidimensional, que requer a adoção de abordagens metodológicas diversificadas. Nenhum método, isoladamente, é capaz de captar a totalidade dos aspetos inerentes à QIF, pelo que a literatura e a prática empresarial têm valorizado cada vez mais a combinação entre abordagens quantitativas e qualitativas (Licerán-Gutiérrez & Cano-Rodríguez, 2019).

Os métodos quantitativos, caracterizados pela objetividade e pela utilização de métricas numéricas, assentam na análise de dados financeiros através de indicadores estatísticos. Entre os mais utilizados destacam-se os modelos baseados em *accruals*, como o de Dechow & Dichev e o Modelo de Jones, que associam elevados níveis de ajustamentos contabilísticos a uma menor fiabilidade da informação. Outros exemplos incluem a análise da persistência dos resultados (partindo do pressuposto de que lucros de elevada qualidade são consistentes e previsíveis ao longo do tempo), a avaliação da relevância valorativa da informação nos preços das ações e a deteção de práticas de gestão de resultados (*earnings management*), frequentemente indicativas de manipulação e conseqüente redução da qualidade informativa (Dănescu & Stejerean (2022); Jones, 1991; Licerán-Gutiérrez & Cano-Rodríguez, 2019).

A Lei de Benford é igualmente aplicada como ferramenta forense de deteção preliminar de anomalias, complementando o Modelo de Jones, que aprofunda a análise dos *accruals*, distinguindo ajustamentos legítimos de manipulações. Esta utilização articulada de diferentes métodos permite reforçar a fiabilidade dos resultados obtidos (Jones, 1991; Silva et al., 2017).

Em paralelo, os métodos qualitativos centram-se na análise de elementos não numéricos, como a transparência da comunicação, a estrutura organizacional e os padrões éticos observados. Tais abordagens recorrem a técnicas como a análise de conteúdo dos relatórios, a avaliação da qualidade das auditorias (especialmente as realizadas por entidades de reconhecida reputação) e a realização de entrevistas com gestores ou outras partes interessadas, de modo a captar perceções subjetivas sobre

a credibilidade da informação financeira. Diferenciam-se, assim, pela profundidade analítica, pela adaptabilidade e pela valorização dos contextos sociais e culturais, ainda que enfrentem limitações relacionadas com a subjetividade e a menor capacidade de generalização (Licerán-Gutiérrez & Cano-Rodríguez, 2019).

A conjugação destas metodologias, conhecida como metodologia mista, representa uma estratégia eficaz para colmatar as fragilidades individuais de cada abordagem, unindo a precisão dos indicadores mensuráveis à riqueza interpretativa da análise qualitativa. A utilização simultânea destas duas abordagens, ao conjugar factos mensuráveis com a análise do contexto, permite uma abordagem mais completa, favorecendo a robustez das evidências, a adequação dos instrumentos de análise ao cenário específico e a descoberta de novas perspetivas de investigação (Licerán-Gutiérrez & Cano-Rodríguez, 2019).

Na prática empresarial, esta complementaridade metodológica revela-se particularmente útil em diversos domínios: na tomada de decisões estratégicas, ao conjugar dados financeiros com perceções de clientes e colaboradores; no desenvolvimento de produtos, através da articulação de inquéritos com grupos focais; e na gestão da mudança, monitorizando impactos e reações humanas.

A relevância desta abordagem é destacada pela sua aplicabilidade em auditorias e pela crescente preocupação com a integridade financeira, especialmente em cenários pós-crise financeira, onde a confiança nos relatórios financeiros é constantemente posta à prova (Costa & Soares, 2022; Nuhu et al., 2024). Estudos recentes (e.g. Martins, 2023) confirmam que a utilização conjunta de diferentes técnicas pode proporcionar uma solução mais eficaz para a identificação de irregularidades, oferecendo maior segurança e credibilidade no ambiente corporativo e reforçando os mecanismos de prevenção de fraudes empresariais.

Para além disso, a literatura tem demonstrado que a QIF influencia diretamente os preços dos ativos, uma vez que lucros de elevada qualidade, por serem mais persistentes e previsíveis, servem de base para estimar fluxos futuros de caixa e dividendos, reduzindo riscos de avaliação e fortalecendo a credibilidade patrimonial (Souza et al., 2024). Neste domínio, destacam-se dois métodos de mensuração: o da persistência dos lucros, estimada a partir do fluxo de caixa e do coeficiente beta da regressão, indicador da sustentabilidade dos resultados ao longo do tempo e o da relevância do preço das ações, avaliado pelo modelo de Ohlson (1995), que relaciona preço de mercado, lucro por ação e valor contabilístico por ação (Souza et al., 2024).

Constata-se, ainda, que relatórios financeiros de baixa qualidade estão associados a menores retornos futuros e a um conteúdo informativo menos útil, levando os investidores a recorrer a avaliações individualistas, crenças divergentes e fontes privadas de informação, o que acentua a incerteza e fomenta divergências no mercado acionista (Silva & Cerqueira, 2021). A este risco acresce a

possibilidade de distorções nas DF e de influência implícita nas percepções dos utilizadores finais, que baseiam as suas decisões de investimento na informação divulgada (Dănescu et al., 2021). Assim, a qualidade da informação deve ser entendida não apenas em termos técnicos, mas também à luz da ética, da transparência e da imparcialidade da gestão, fatores que reforçam a necessidade de uma avaliação sustentada em diferentes métodos.

Segundo Gillette e Stinson (2022), as empresas com maior credibilidade em relação aos seus resultados possuem mais margem para adotar comportamentos questionáveis, nomeadamente nas suas estratégias de minimização de impostos.

Deste modo, a integração de abordagens quantitativas e qualitativas revela-se essencial, tanto para a investigação académica como para a prática empresarial, assegurando maior fiabilidade dos relatórios financeiros, fortalecendo os mecanismos de prevenção de fraudes e promovendo um ambiente corporativo mais transparente, ético e credível, capaz de aumentar a confiança dos investidores e dos restantes *stakeholders* na informação divulgada pelas empresas.

Os principais métodos aplicados na avaliação da QIF envolvidos neste estudo são a Lei de Benford e o Modelo de Jones.

A Lei de Benford contribui para a identificação de irregularidade e fraudes, e pode ser utilizada de forma complementar ao Modelo de Jones, que visa analisar a manipulação de resultados, ou seja, a Lei de Benford verifica os dados iniciais e o Modelo de Jones aprofunda essa análise relativamente à qualidade dos *accruals*, distinguindo os ajustes legítimos dos manipulados (Jones, 1991; Silva et al., 2017).

A utilização combinada da Lei de Benford e do Modelo de Jones permite uma análise mais rigorosa: enquanto a Lei de Benford serve como uma ferramenta inicial de deteção de anomalias, o Modelo de Jones aprofunda essa análise, focando na manipulação dos resultados. Esta abordagem integrada é essencial para reforçar a fiabilidade dos relatórios financeiros e prevenir fraudes empresariais. Assim, a conjugação de metodologias quantitativas e qualitativas, com o recurso a ferramentas estatísticas e técnicas forenses, oferece uma abordagem eficaz para avaliar a QIF, beneficiando tanto a investigação académica quanto a prática empresarial, e fortalecendo a confiança dos *stakeholders* na informação financeira divulgada pelas empresas.

2. Metodologia de Investigação

2.1. Objetivo do Estudo e Hipóteses de Investigação

O objetivo principal deste estudo é avaliar a eficácia da combinação da Lei de Benford com o modelo baseado nos *accruals* discricionários - Modelo Jones (1991) - para aferir a QIF e identificar indícios de possível manipulação dos resultados de empresas do setor da restauração e similares e o impacto que isto tem na minimização do pagamento de impostos ao Estado. Para o efeito, será explorada e explicada a aplicabilidade da Lei de Benford e do modelo de *accruals* discricionários de Jones (1991) na deteção de anomalias nos registos financeiros das empresas em análise.

De forma a aferir a relevância da combinação destes dois métodos será realizada uma comparação entre a eficácia individual de cada um e a eficácia conjunta no que respeita à identificação de indícios de manipulação dos resultados com o objetivo de minimizar a carga tributária. Relativamente aos incentivos fiscais, estes assumem um papel importante no estímulo ao crescimento económico e na atração de investimentos em Portugal. O Código Fiscal do Investimento proporciona às empresas oportunidades significativas de redução de encargos fiscais e otimização da eficiência tributária (Amorim et al., 2020; Decreto-Lei n.º 162/2014, 2014).

No entanto, de acordo com Silva et al. (2022), a não conformidade fiscal continua a ser um desafio. Por conseguinte, é crucial compreender as motivações subjacentes à conformidade fiscal, para se elaborarem estratégias eficazes que visem a minimização da evasão fiscal.

Segundo Marques et al. (2016), os incentivos fiscais atuam como instrumentos importantes para o planeamento tributário das empresas, permitindo-lhes diminuir as suas obrigações fiscais e ao mesmo tempo contribuir para o desenvolvimento económico. Estes autores concluíram que os incentivos fiscais podem reduzir entre 30% a 40% os impostos a pagar ao Estado. Contudo, referem também que a consciencialização destes incentivos por parte da área de contabilidade é ainda insuficiente, evidenciando a necessidade de melhorar a comunicação sobre os benefícios fiscais (Marques et al., 2016).

Face ao exposto, estabelece-se a seguinte hipótese de investigação (Gillette & Stinson, 2022; Silva et al., 2017):

H1: A combinação da Lei de Benford com o modelo de *accruals* discricionários de Jones (1991) permite identificar uma melhor classificação das empresas com indícios de manipulação de resultados com o incentivo da minimização do pagamento de impostos ao Estado.

A investigação seguirá uma abordagem quantitativa, recorrendo à análise estatística dos dados financeiros recolhidos, de modo a testar empiricamente as hipóteses formuladas.

2.2. Descrição do Instrumento de Recolha de Dados

O instrumento de recolha de dados utilizado neste estudo foi a base de dados Sistema de Análise de Balanços Ibéricos (SABI) que reúne informações financeiras detalhadas de empresas em Portugal e em Espanha. Esta plataforma fornece dados atualizados a partir de fontes oficiais como o Registo Comercial, o Banco de Portugal e outras entidades reguladoras.

Para a presente investigação, foram selecionadas empresas com o Código de Atividade Económica (CAE) 56, correspondente ao setor da restauração e similares, em Portugal. Adicionalmente, foram extraídos dados financeiros essenciais, nomeadamente: ATF, caixa e depósitos bancários, total do ativo, total do ativo corrente, total do passivo corrente, financiamentos obtidos correntes, vendas e serviços prestados, bem como gastos ou reversões de depreciação e de amortização, resultado antes de imposto, imposto sobre o rendimento do período, EBITDA e pagamento ou recebimento do imposto sobre o rendimento. A análise incide sobre o período compreendido entre 2019 e 2024.

2.3. População vs. Amostra

No contexto deste estudo, a população corresponde ao universo de empresas do setor da restauração e similares em Portugal incluídas na base de dados SABI, no período em análise. Esta população é composta por todas as entidades que operam sob o código CAE 56 (Restauração e similares). Assim, a amostra inicial contempla 69.507 empresas e sofreu a seguinte evolução, conforme se apresenta na seguinte Tabela 4:

Tabela 4 - Evolução da amostra selecionada

Descrição	N.º de Obs.
Total de observações (empresas-ano) disponíveis na SABI para o período 2019-2024. Encontravam-se nesta situação 69.507 empresas.	417.042
Total de observações (empresas-ano) após eliminação das empresas que possuíam variáveis sem valor disponível, valores negativos ou nulos (zero) em variáveis onde esses valores não são economicamente plausíveis. Número de observações utilizado para os cálculos de conformidade da Lei de Benford.	127.030
Total de observações (empresas-ano) final. Número de observações utilizadas para a estimativa dos <i>accruals</i> discricionários.	89.219

Fonte: Elaboração própria.

A Dimensão da amostra total e dos $LBC=1$, por dimensão da empresa, encontram-se apresentadas no Apêndice 1. No Apêndice 2 apresenta-se o número de empresas por ano.

2.4. Descrição dos Métodos de Tratamento dos Dados

O tratamento de dados foi realizado com o recurso ao software livre Rstudio, permitindo assim uma análise detalhada e rigorosa. Inicialmente, procedeu-se a uma análise estatística descritiva para caracterizar as variáveis em estudo. Foram calculadas medidas como a média, a mediana, o desvio padrão, o coeficiente de variação para o desvio padrão, o mínimo e o máximo de cada uma das variáveis e o peso de determinadas variáveis em relação ao total (exemplo: a proporção de AFT e de caixa e depósitos à ordem no ativo total). Para garantir a validade dos resultados, foi realizada a verificação da representatividade da amostra, com base em cálculos estatísticos que asseguram a correspondência do seu tamanho da amostra em relação à população.

Posteriormente, foi aplicada a Lei de Benford, um método utilizado para avaliar a distribuição do primeiro dígito significativo em variáveis financeiras. Esta técnica permite identificar desvios em relação à distribuição teórica esperada, o que pode indicar a presença de anomalias ou manipulação de dados. Em complemento, foi implementado o Modelo de Jones, que se baseia na estimação de uma regressão linear, na qual os AT são explicados por variáveis como a variação da receita e o ativo fixo tangível (Propriedade, planta e equipamento - PPE). A diferença entre os *accruals* observados e os *accruals* estimados permite identificar a parcela discricionária, que pode indicar práticas de manipulação dos resultados.

Os resultados foram analisados criticamente, recorrendo a testes estatísticos adequados para validar as hipóteses e avaliar a conformidade dos dados. Este processo permite não apenas compreender as características da amostra, mas também identificar padrões e possíveis desvios que possam indicar irregularidades na informação financeira.

2.4.1. Aplicação da Lei de Benford

A Lei de Benford é um método estatístico utilizado para analisar a frequência dos primeiros dígitos significativos para um conjunto de dados numéricos e é um instrumento relevante para a avaliação da naturalidade e da integridade das informações reportadas. Esta lei estatística, defende que, em muitos conjuntos de dados reais, os valores numéricos tendem a começar por dígitos mais pequenos, sendo o dígito 1 o mais recorrente, ao surgir como algarismo inicial em cerca de 30% das ocorrências, enquanto o número 9 aparece com menor frequência (Silva et al., 2017).

Através da comparação entre a frequência observada dos primeiros dígitos de variáveis como o total do ativo, o volume de negócios ou os montantes em caixa, e a distribuição teórica prevista pela Lei de Benford, é possível identificar discrepâncias relevantes que possam sinalizar incoerências, erros de registo ou até possíveis tentativas de manipulação (Silva et al., 2017). Embora esta regularidade não se aplique a todas as variáveis financeiras, a sua utilização como metodologia complementar revela-se valiosa, sobretudo em procedimentos de auditoria ou em processos de validação da consistência dos dados contabilísticos apresentados pelas entidades analisadas (Antunes et al., 2023; Silva et al., 2017).

Importa referir que a Lei de Benford não é universalmente aplicável a todas as variáveis financeiras, uma vez que a sua validade depende de certas características estatísticas dos dados. Em particular, esta lei tende a verificar-se em conjuntos de dados que se distribuem de forma não uniforme, abrangem várias ordens de grandeza e não estão sujeitos a limites arbitrários (Antunes et al., 2023).

A Lei de Benford pode ser avaliada através de diferentes procedimentos estatísticos. O teste do primeiro dígito, apresentado por Nigrini e Mittermaier (1997), tem como objetivo principal verificar a conformidade dos dados com a distribuição teórica da referida Lei, não se destinando diretamente à identificação de fraude. Contudo, quando os resultados revelam desvios expressivos em relação ao padrão esperado, estes podem constituir indícios de potenciais irregularidades, recomendando a realização de análises adicionais.

Posteriormente, Nigrini (2011) salientou algumas limitações deste teste, nomeadamente em conjuntos de dados de grande dimensão, nos quais é possível observar uma conformidade aparente com a Lei de Benford, ainda que existam anomalias ou inconsistências subjacentes. Neste contexto, o autor destaca o teste dos dois primeiros dígitos como o mais adequado, dado que este permite captar discrepâncias que não são identificadas pelos testes do primeiro ou do segundo dígito isoladamente, revelando-se, assim, fundamental na análise de bases de dados extensas.

Para aplicar a Lei de Benford aos dados recolhidos, é extraído o primeiro dígito significativo de variáveis financeiras relevantes constantes na base de dados, nomeadamente as vendas e prestações de serviços. De seguida, procede-se ao cálculo da frequência observada de cada dígito inicial (de 1 a 9). Assim, a comparação entre a distribuição observada e a distribuição teórica permite avaliar a conformidade dos dados com a Lei de Benford, identificando eventuais desvios que possam indiciar anomalias ou manipulação nos registos contabilísticos. Importa salientar que este procedimento não se destina diretamente à identificação de fraude.

Subsequentemente, procede-se, na variável as vendas e prestações de serviço, à extração dos dois primeiros dígitos (de 10 a 99), sendo igualmente efetuada a comparação entre a distribuição observada e a distribuição teórica. Esta análise permite aferir a conformidade dos dados com padrões esperados em contextos financeiros naturais e, potencialmente, identificar desvios que possam indicar inconsistências ou irregularidades nos registos contabilísticos.

Desta forma, após a comparação entre a distribuição observada e a teórica prevista pela Lei de Benford, procede-se à aplicação de testes estatísticos, nomeadamente o teste do qui-quadrado e o teste z, com o objetivo de detetar eventuais irregularidades nos dados. Para tal, formulam-se as seguintes hipóteses estatísticas:

H0: Não existem diferenças estatisticamente significativas entre a probabilidade observada e a probabilidade teórica (ou seja, os dados seguem a distribuição prevista pela Lei de Benford).

H1: Existem diferenças estatisticamente significativas entre a probabilidade observada e a probabilidade teórica (ou seja, os dados não seguem a distribuição prevista pela Lei de Benford).

A análise foi realizada com o auxílio do software livre Rstudio e do Microsoft Excel, os quais possibilitaram a automatizar o processo de extração dos dígitos, efetuar os cálculos estatísticos e interpretar os resultados de forma rigorosa e detalhada.

2.4.2. Estimativa dos *Accruals* Discricionários

A estimativa dos *accruals* discricionários foi realizada com recurso ao modelo de Jones (1991), amplamente utilizado na literatura para identificar potenciais práticas de manipulação de resultados. Este modelo parte da decomposição dos AT em componentes discricionários e não discricionários, sendo estes últimos explicados por variáveis que refletem a atividade económica da empresa.

Para aplicar o modelo à base de dados disponível, procedeu-se em primeiro lugar ao cálculo dos AT, definidos como a diferença entre o resultado líquido e os fluxos de caixa operacionais, representados pela fórmula seguinte:

$$AT = (\Delta AC - \Delta Cx) - (\Delta PC - \Delta Dív) - Dep \quad [7]$$

Nota: ΔAC - Variação do ativo corrente; ΔCx - Variação do caixa ou equivalentes a caixa; ΔPC - Variação do passivo corrente; $\Delta Dív$ - Variação da dívida a curto prazo; Dep - Despesa com depreciação e amortização. Fonte: Elaboração própria, com recurso a Cupertino e Martinez (2008).

Na ausência destes fluxos, os AT podem ser estimados indiretamente com base em variações do capital circulante e ajustamentos ao AFT e à depreciação.

Posteriormente, foi estimada uma regressão linear para cada ano seguido em cada empresa, em que os AT normalizados pelo total do ativo do período anterior constituem a variável dependente, sendo as variáveis independentes: (i) o inverso do total do ativo do ano anterior, (ii) a variação dos rendimentos de vendas e prestação de serviços, e (iii) o montante líquido dos AFT, todas igualmente normalizadas.

Os resíduos da regressão representam os *accruals* discricionários, que são objeto de análise com vista à deteção de comportamentos anómalos ou práticas de manipulação contabilística. Desta forma, estes são representados pela fórmula seguinte (Costa e Soares, 2022):

$$\left(\frac{TA_{it}}{A_{it-1}}\right) = \alpha_0 + \alpha_1 \left(\frac{1}{A_{it-1}}\right) + \beta_1 \left(\frac{\Delta R_{it}}{A_{it-1}}\right) + \beta_2 \left(\frac{AFT_{it}}{A_{it-1}}\right) + \epsilon_{it} \quad [8]$$

Nota: TA_{it} - Acréscimos totais no ano t para a empresa i; ΔR_{it} - Receitas no ano t menos receitas no ano t-1 para a empresa i; AFT_{it} - Ativos Fixos brutos, instalações e equipamentos no ano t para a empresa i; A_{it-1} - Ativos totais no ano t-1 para a empresa i; ϵ_{it} - Termo de erro no ano t para a empresa i; β_0 - Constante; $\beta_1, \beta_2, \beta_3$ - Coeficientes das covariáveis. Fonte: Elaboração própria, com recurso a Costa e Soares, (2022) e Jones (1991).

Após a aplicação da Lei de Benford e a identificação das combinações críticas (isto é, aquelas em que o valor de p se revelou inferior a 0,05, indicando desvios estatisticamente significativos), procedeu-se à criação de uma variável *dummy*, designada Lei de Benford Crítica (LBC). Nesta variável, as observações correspondentes às combinações críticas assumiram o valor 1, enquanto as restantes assumiram o valor 0, permitindo distinguir, de forma binária, as entidades cujos dados apresentavam discrepâncias relevantes face à distribuição teórica esperada.

Posteriormente, foi estimado o modelo de Jones, com o objetivo de apurar os *accruals* discricionários (DAC), indicadores associados a potenciais práticas de manipulação de resultados. De seguida, procedeu-se ao cálculo do valor absoluto dos DAC, originando a variável DACA, a qual traduz a magnitude das variações discricionárias, independentemente da sua direção (positiva ou negativa).

Com base nestes resultados, criou-se a variável NDIM, que reflete a dimensão das empresas da amostra, através de uma escala ordinal em que: o valor 1 corresponde a microempresas, o valor 2 a pequenas empresas, o valor 3 a médias empresas, e o valor 4 a grandes empresas.

Após a definição e estatísticas descritivas, procedeu-se à análise das correlações entre a dimensão empresarial (NDIM), DACA, os Impostos sobre o Rendimento do Período (ISRI) e a LBC. Este procedimento teve como objetivo identificar potenciais relações entre a dimensão da entidade, o nível de manipulação de resultados e a carga fiscal, permitindo uma compreensão mais aprofundada das dinâmicas financeiras e contabilísticas que caracterizam as entidades pertencentes ao setor da restauração e similares.

3. Apresentação e Análise dos Resultados Empíricos

3.1. Caracterização da Amostra

A base de dados, como explicado anteriormente, foi sujeita a uma limpeza, a qual procurou garantir a consistência e a fiabilidade dos dados analisados, eliminando registos que poderiam comprometer a robustez estatística da análise. Embora se trate de uma população finita, o seu tamanho exato é desconhecido e não relevante para o presente estudo. Assim, não se apresenta uma percentagem de margem de erro, considerando-se, em vez disso, as 89 219 observações disponíveis na base de dados como base para a análise, as quais são consideradas suficientemente grandes para permitir inferências válidas e seguras sobre o universo em estudo.

Com o objetivo de caracterizar a amostra e compreender o comportamento das variáveis analisadas, a Tabela 5 apresenta as estatísticas descritivas das variáveis mais significativas. São incluídas medidas de tendência central (média e mediana) e de dispersão (desvio padrão, valores mínimo e máximo), bem como o coeficiente de variação, permitindo uma leitura abrangente da distribuição e da variabilidade dos dados. As restantes variáveis utilizadas na análise encontram-se apresentadas no Apêndice 3, de forma a complementar a informação aqui resumida.

Tabela 5 - Tabela de estatística descritiva

Descrição	Ano	A	AFT	TPC	FOC	VPS	GRDA	RAI	ISRI
Média	2020	245 227,30 €	96 995,12 €	97 616,51 €	20 124,69 €	236 625,10 €	12 323,08 €	-16 127,43 €	1 150,45 €
	2021	268 113,90 €	100 498,30 €	109 540,80 €	20 376,34 €	284 932,60 €	12 606,18 €	6 514,45 €	3 107,44 €
	2022	293 696,40 €	107 465,40 €	122 632,10 €	19 793,36 €	431 521,80 €	13 528,91 €	23 198,19 €	5 128,44 €
	2023	324 798,10 €	119 461,70 €	131 048,20 €	20 174,40 €	502 237,70 €	15 277,94 €	28 153,36 €	6 477,03 €
	2024	1 173 481,00 €	410 962,90 €	684 515,90 €	0,00 €	3 016 646,00 €	101 927,84 €	26 536,25 €	5 830,89 €
Desvio Padrão	2020	2 158 958,00 €	1 293 289,00 €	633 481,90 €	283 801,60 €	1 608 210,00 €	151 227,23 €	192 505,18 €	41 615,27 €
	2021	2 283 549,00 €	1 369 946,00 €	686 078,30 €	203 017,40 €	1 876 159,00 €	152 162,87 €	258 760,31 €	66 378,13 €
	2022	2 465 938,00 €	1 456 701,00 €	763 717,30 €	148 116,60 €	2 559 169,00 €	161 871,33 €	389 986,01 €	103 638,44 €
	2023	3 560 729,00 €	1 723 106,00 €	868 592,60 €	206 455,60 €	3 151 735,00 €	230 131,18 €	548 777,41 €	162 061,15 €
	2024	1 583 894,00 €	517 981,30 €	898 880,30 €	0,00 €	4 168 160,00 €	137 541,32 €	35 949,33 €	7 579,58 €
Mediana	2020	66 135,62 €	12 874,18 €	29 369,97 €	0,00 €	95 663,68 €	2 180,49 €	-9 302,58 €	0,00 €
	2021	72 774,09 €	13 442,76 €	34 269,70 €	0,00 €	114 364,40 €	2 277,71 €	-301,84 €	132,53 €
	2022	80 307,30 €	14 121,94 €	39 073,69 €	0,00 €	173 765,30 €	2 469,99 €	2 226,46 €	332,28 €
	2023	86 342,90 €	15 726,08 €	42 226,01 €	0,00 €	202 820,70 €	2 792,98 €	2 687,58 €	464,72 €
	2024	1 173 481,00 €	410 962,90 €	684 515,90 €	0,00 €	3 016 646,00 €	101 927,84 €	26 536,25 €	5 830,89 €
Mínimo	2020	4,88 €	0,00 €	0,00 €	0,00 €	1,63 €	-35 490,79 €	-6 950 486,65 €	-2 063 357,27 €
	2021	8,48 €	0,00 €	0,00 €	0,00 €	4,40 €	-14 343,33 €	-4 993 023,11 €	-1 326 208,31 €
	2022	1,71 €	0,00 €	0,00 €	0,00 €	2,15 €	-1 454,72 €	-5 108 887,41 €	-5 428 117,17 €
	2023	4,94 €	0,00 €	0,00 €	0,00 €	1,88 €	-21 065,94 €	-7 853 006,91 €	-5 772 701,78 €
	2024	53 498,54 €	44 694,79 €	48 911,54 €	0,00 €	69 312,30 €	4 671,44 €	1 116,23 €	471,32 €

Descrição	Ano	A	AFT	TPC	FOC	VPS	GRDA	RAI	ISRI
Máximo	2020	190 580 900,00 €	137 322 000,00 €	50 801 900,00 €	31 404 160,00 €	102 840 000,00 €	13 433 797,69 €	19 869 920,64 €	4 854 026,54 €
	2021	200 924 700,00 €	148 882 000,00 €	44 117 390,00 €	19 385 900,00 €	112 282 100,00 €	13 595 662,48 €	31 348 945,30 €	8 697 526,20 €
	2022	216 166 100,00 €	159 890 900,00 €	50 489 770,00 €	11 808 300,00 €	143 751 300,00 €	14 983 551,94 €	46 949 729,97 €	12 957 576,81 €
	2023	355 019 600,00 €	185 635 000,00 €	53 846 910,00 €	18 184 670,00 €	199 118 300,00 €	27 341 389,85 €	65 331 558,49 €	19 620 588,06 €
	2024	2 293 463,00 €	777 231,00 €	1 320 120,00 €	0,00 €	5 963 980,00 €	199 184,24 €	51 956,26 €	11 190,46 €
Coefficiente de Variação	2020	880,4%	1333,4%	648,9%	1410,2%	679,6%	1227,2%	-1193,7%	3617,3%
	2021	851,7%	1363,2%	626,3%	996,3%	658,5%	1207,0%	3972,1%	2136,1%
	2022	839,6%	1355,5%	622,8%	748,3%	593,1%	1196,5%	1681,1%	2020,9%
	2023	1096,3%	1442,4%	662,8%	1023,4%	627,5%	1506,3%	1949,2%	2502,1%
	2024	135,0%	126,0%	131,3%	NA	138,2%	134,9%	135,5%	130,0%

Nota: A – Total do ativo; AFT – Ativo fixo tangível; TPC – Total do passivo corrente; FOC – Financiamentos obtidos correntes; VPS – Vendas e serviços prestados; GRDA – Gastos e reversões de depreciação e de amortização; RAI – Resultado antes de imposto; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria, com recurso aos softwares Rstudio e ao Excel.

A Tabela 5 apresenta as estatísticas descritivas das principais variáveis financeiras do conjunto de empresas analisadas, relativas ao período de 2020 a 2024. O ano de 2019 não foi incluído na análise, uma vez que, para a estimação do Modelo de Jones, foi necessário eliminar esse período devido ao cálculo das variações anuais. A análise das variáveis selecionadas permite compreender não apenas a evolução média dos principais elementos patrimoniais, do volume de negócios e das depreciações, mas também a sua dispersão e variabilidade, evidenciando o grau de heterogeneidade existente entre as empresas da amostra. Os dados são apresentados com base em medidas estatísticas como a média, desvio padrão, mediana, valores máximos e mínimos, além do coeficiente de variação, permitindo uma leitura abrangente da evolução e da dispersão dos resultados no setor. No ano de 2024, conforme se observa na Tabela 5, a média e a mediana coincidem. Tal facto deve-se à existência de apenas duas empresas nas observações desse ano.

O total do ativo evidencia um crescimento expressivo ao longo do período analisado, passando de uma média de 245.227,30€ em 2020 para 1.173.481,00€ em 2024, o que traduz um reforço da base patrimonial das empresas. Tendências semelhantes são observadas em AFT, TPC e VPS, o que sugere um aumento global do volume financeiro ou da dimensão das operações consideradas. Em contraste, a variável FOC apresenta um valor nulo em 2024, o que poderá indicar ausência de registos.

O desvio padrão e o coeficiente de variação revelam, contudo, que esta evolução foi acompanhada por uma grande variabilidade nos dados. Entre 2020 e 2023, os desvios padrões atingem valores muito elevados, demonstrando fortes flutuações e possíveis assimetrias nas distribuições. Em conjunto com mínimos próximos de zero, máximos muito elevados e médias sistematicamente acima das medianas, estes factos confirmam heterogeneidade acentuada entre empresas e uma assimetria positiva marcada.

Tendo em conta que, em 2024, a amostra se encontra enviesada para grandes empresas, que tendem a divulgar as suas contas mais cedo do que as PME, esse ano poderá desvirtuar a comparabilidade com os restantes períodos analisados. Por essa razão, não é atribuída especial ênfase à análise de 2024, sendo as principais conclusões reportadas ao período de 2020 a 2023.

O coeficiente de variação, por sua vez, ultrapassa frequentemente os 100%, o que indica uma dispersão superior à média e evidencia a presença de valores extremos. Este facto reflete a forte heterogeneidade do setor, sugerindo que a média é fortemente influenciada por um número restrito de empresas de grande dimensão. A diferença entre a média e a mediana reforça esta interpretação, apontando para uma distribuição assimétrica à direita, onde poucas entidades concentram valores muito superiores à maioria.

Em 2024, observa-se uma redução generalizada do desvio padrão e do coeficiente de variação, o que poderá traduzir uma maior estabilidade e homogeneidade dos resultados financeiros. Contudo, esta diminuição também pode indicar ausência de registos ou a limitação na recolha de determinados dados nesse período.

A mediana acompanha a tendência de crescimento da média, embora apresentem valores geralmente inferiores. Esta diferença confirma a existência de assimetria positiva, ou seja, a presença de alguns valores muito elevados que influenciam as médias para cima.

A análise dos valores mínimos e máximos reforça a perceção de uma distribuição bastante ampla dos dados. Enquanto os mínimos se mantêm baixos ou mesmo nulos em várias variáveis, os máximos apresentam aumentos acentuados, sobretudo no total do ativo, AFT e GRDA, o que demonstra a existência de valores atípicos elevados. Verifica-se um alargamento progressivo da amplitude entre os valores mínimos e máximos ao longo do período em análise, indicando que, embora se observe uma melhoria do desempenho médio, continuam a existir disparidades claras entre as diferentes observações.

De forma global, o período compreendido entre 2020 e 2023 caracteriza-se por forte crescimento, mas também elevada volatilidade. Já em 2024, nota-se uma tendência de estabilização aparente, com diminuição relativa da dispersão e possível consolidação dos resultados financeiros. Esta evolução pode refletir um amadurecimento dos processos operacionais ou uma gestão mais eficiente dos recursos, dado o reduzido número de observações.

Em conclusão, a análise permite identificar uma trajetória de crescimento consistente ao longo do período em estudo, embora acompanhada por uma acentuada flutuação nos primeiros anos. Entre os anos de 2020 a 2023, o crescimento foi desigual e concentrado nas empresas de maior dimensão, com os dados a revelarem forte heterogeneidade e assimetria positiva. A redução observada no coeficiente de variação em 2024 poderá, numa primeira leitura, ser interpretada como um indício de maior

consistência e previsibilidade dos resultados. Todavia, esta evidência deve ser analisada com cautela, uma vez que o número reduzido de observações registrado nesse ano poderá ter condicionado de forma significativa os indicadores estatísticos, conduzindo a uma atenuação artificial da dispersão dos valores. Assim, a aparente estabilização verificada em 2024 poderá refletir mais uma limitação de natureza amostral, do que uma efetiva consolidação da situação financeira das entidades analisadas.

Adicionalmente, a presença de valores nulos e de variáveis com comportamentos irregulares, nomeadamente FOC e RAI, reforça a necessidade de proceder a uma avaliação crítica da metodologia aplicada e da qualidade dos dados recolhidos, de forma a assegurar a robustez e fiabilidade das conclusões obtidas.

3.2. Análise dos Resultados da Aplicação da Lei de Benford

A aplicação da Lei de Benford tem vindo a revelar-se uma ferramenta estatística relevante na deteção de anomalias, nomeadamente em contextos de auditoria, deteção de fraude e validação de integridade de dados (Silva et al., 2017).

No presente estudo, procedeu-se à aplicação da referida lei sobre a variável vendas e prestações de serviços, para a amostra de 127.030, com o intuito de avaliar o grau de conformidade entre a distribuição observada dos primeiros e segundos dígitos e a distribuição teórica proposta pela Lei de Benford.

A análise contempla a comparação entre as frequências observadas e teóricas, bem como a utilização do teste de qui-quadrado e o teste z como método de aferição estatística da significância das divergências detetadas.

Os resultados obtidos são apresentados seguidamente, acompanhados de uma reflexão crítica sobre os desvios identificados e as suas possíveis implicações.

Inicialmente, começar-se-á por extrair o primeiro dígito, e continuamente, o primeiro e o segundo dígito da variável vendas e prestações de serviços da amostra. De seguida, fez-se a contagem desses dígitos e procedeu-se ao cálculo das frequências observadas para o primeiro dígito, como se pode verificar na Tabela 6.

Tabela 6 - Frequências do primeiro dígito

Números	Contagem dos números	Frequência observada	Frequência teórica	Desvio	P-Value teste Z	P-Value teste χ^2	Nível de significância
1	39332	30,96%	30,10%	0,009	0,000	0,000	<0,05
2	22634	17,82%	17,61%	0,002	0,051	0,076	>=0,05
3	15579	12,26%	12,49%	0,002	0,015	0,023	<0,05
4	11911	9,38%	9,69%	0,003	0,000	0,000	<0,05
5	9705	7,64%	7,92%	0,003	0,000	0,000	<0,05
6	8278	6,52%	6,70%	0,002	0,009	0,012	<0,05
7	7288	5,74%	5,80%	0,001	0,338	0,353	>=0,05
8	6469	5,09%	5,12%	0,000	0,657	0,665	>=0,05
9	5834	4,59%	4,58%	0,000	0,830	0,834	>=0,05

Fonte: Elaboração própria, com recurso ao Excel.

De forma geral, as frequências observadas são bastante próximas das frequências teóricas, com desvios reduzidos (entre 0,000 e 0,009), o que indica um elevado grau de conformidade global com a Lei de Benford. No entanto, a análise estatística detalhada permite identificar algumas discrepâncias significativas em determinados dígitos.

Por um lado, nos dígitos 1, 3, 4, 5 e 6, os valores de *p-value* dos testes Z e χ^2 são inferiores a 0,05, o que implica a rejeição da hipótese nula (H_0), para esses casos, evidência diferenças estatisticamente relevantes entre as frequências observadas e as esperadas. Por outro lado, para os dígitos 2, 7, 8 e 9, os valores de p são iguais ou superiores a 0,05, não havendo motivo para rejeitar a hipótese nula, o que demonstra que, nesses casos, a distribuição observada não difere significativamente da teórica.

Assim, os resultados indicam que a amostra segue parcialmente a Lei de Benford, mantendo a tendência esperada, (maior incidência do dígito 1 e redução gradual até ao dígito 9). Estes desvios podem dever-se a características específicas do conjunto de dados, como o tipo de transações analisadas, o período temporal, ou o arredondamento de valores, fatores que podem afetar o comportamento do primeiro dígito.

Em síntese, a análise empírica permite concluir que, de forma geral, a distribuição dos primeiros dígitos apresenta conformidade global com a Lei de Benford, ainda que essa correspondência não seja integral. A rejeição da hipótese nula em alguns casos evidencia inconformidades pontuais, mas não compromete a aderência global da amostra ao padrão teórico esperado. Deste modo, considera-se que os dados

analisados seguem, em termos gerais, a distribuição de Benford, embora se verifiquem diferenças estatisticamente significativas em determinados dígitos. Estas discrepâncias justificam a realização de uma análise complementar, com vista a identificar as potenciais causas subjacentes.

Seguidamente à análise do primeiro dígito, efetuou-se a mesma análise para os dois primeiros dígitos, como se pode observar no Apêndice 4 a tabela completa. Na tabela abaixo apresentam-se apenas os casos em que dois primeiros dígitos assumem valores críticos, sendo estes os valores que dão origem à variável LBC.

Tabela 7 - Frequências do primeiro e segundo dígito que tomam valores críticos

Números	Contagem dos números	Frequência observada	Frequência teórica	Desvio	P-Value teste Z	P-Value teste χ^2
12	4593	3,62%	3,48%	0,001	0,006	0,007
15	3789	2,98%	2,80%	0,002	0,000	0,000
18	3126	2,46%	2,35%	0,001	0,008	0,009
44	1150	0,91%	0,98%	-0,001	0,011	0,011
46	1107	0,87%	0,93%	-0,001	0,021	0,021
49	1042	0,82%	0,88%	-0,001	0,029	0,030
51	974	0,77%	0,84%	-0,001	0,003	0,003
59	856	0,67%	0,73%	-0,001	0,019	0,020
60	1007	0,79%	0,72%	0,001	0,002	0,002
62	800	0,63%	0,69%	-0,001	0,005	0,005
64	794	0,63%	0,67%	0,000	0,036	0,036
80	760	0,60%	0,54%	0,001	0,004	0,004
84	715	0,56%	0,51%	0,000	0,015	0,015
85	593	0,47%	0,51%	0,000	0,040	0,040
87	578	0,46%	0,50%	0,000	0,036	0,037

Fonte: Elaboração própria, com recurso ao Excel.

De forma geral, as frequências observadas revelam-se muito próximas das frequências teóricas, com desvios residuais reduzidos (inferiores na maioria a 0,001), o que indica uma forte conformidade global com a Lei de Benford na análise conjunta do primeiro e segundo dígito.

Ainda assim, observam-se algumas combinações críticas, nomeadamente 12, 15, 18, 44, 46, 49, 51, 59, 60, 62, 64, 80, 84, 85 e 87, nas quais os valores de p são inferiores a 0,05, conduzindo à rejeição

da hipótese nula (H_0). Estas discrepâncias, embora pontuais, indicam diferenças estatisticamente significativas entre as frequências observadas e as teóricas, podendo refletir características específicas da amostra, como o tipo de transações, o período temporal ou práticas de arredondamento que afetem a distribuição dos dígitos.

Tendo em conta que os dados analisados provêm do setor da restauração e similares, tais variações poderão igualmente resultar de padrões específicos de faturação, de valores frequentemente arredondados, de estruturas de preços pré-definidas em menus ou sistemas de ponto de venda, bem como de políticas comerciais padronizadas que tendem a limitar a variabilidade dos montantes registados.

Adicionalmente, a predominância de transações em numerário, a utilização de sistemas automatizados de registo e as diferenças nos volumes e tipologia das operações entre estabelecimentos podem contribuir para a introdução de assimetrias na distribuição dos valores monetários observados, explicando, assim, desvios pontuais em relação à distribuição teórica esperada. Assim, os desvios identificados não configuram, por si só, indícios de anomalias, mas refletem antes as dinâmicas e constrangimentos próprios do funcionamento do setor.

A maioria dos pares de dígitos, contudo, apresenta valores de p superiores a 0,05, não havendo motivo para rejeitar H_0 . Assim, conclui-se que os dados seguem, de forma global, a distribuição prevista pela Lei de Benford de dois dígitos, apresentando apenas inconformidades localizadas e de reduzida magnitude, as quais não comprometem a regularidade e coerência estatística do conjunto de dados, mas justificam uma verificação complementar das causas subjacentes.

3.3. Análise dos Resultados do Modelo Baseado nos *Accruals*

A presente subsecção tem como objetivo analisar e interpretar os resultados obtidos a partir da estimação do modelo de Jones (1991), aplicado às empresas do setor da restauração e similares. Este modelo é amplamente reconhecido na literatura como uma ferramenta robusta para detetar manipulação de resultados, ao permitir a decomposição dos AT em componentes discricionários e não discricionários.

Através da aplicação da regressão linear, procurou-se estimar a componente normal dos *accruals*, explicada por fatores económicos e operacionais, e, por diferença, identificar a componente discricionária, potencialmente associada a ajustamentos contabilísticos de natureza subjetiva ou intencional. Os resíduos da regressão, representativos dos DAC, funcionam como um *proxy* (uma medida indireta) da parte dos *accruals* associada à discricionariedade contabilística, que pode resultar de decisões de gestão ou de julgamentos profissionais inerentes ao processo contabilístico.

O modelo Jones apresenta bom ajuste estatístico: $R^2 = 0,385$ (ajustado igual) e $F = 18\,634,7$ com $p < 0,0001$. O coeficiente de variação é muito elevado em valor absoluto ($\approx 4\,939\%$), o que indica um desfasamento de escala entre a variabilidade dos resíduos e a média da variável dependente. Mantendo a especificação, conclui-se que o modelo é globalmente significativo, explica cerca de 38,5% da variação dos *accruals* e evidencia um papel dominante do termo de escala (INVA), com coeficiente muito grande e negativo (estatística $t = -203,83$; $p < 0,0001$), uma relação negativa entre a variação das receitas (VEPSD) e os *accruals* (estatística $t = -105,54$; $p < 0,0001$), e um efeito modesto do ativo fixo tangível (AFTD), ainda que estatisticamente significativo (estatística $t = -2,79$; $p = 0,0052$).

Após a aplicação da regressão linear às variáveis em análise, obteve-se a seguinte expressão linear geral:

$$\left(\frac{\text{TAit}}{A_{it-1}}\right) = 0,34497 - 8825,32 \left(\frac{1}{A_{it-1}}\right) - 0,11804 \left(\frac{\Delta R_{it}}{A_{it-1}}\right) - 0,20473 \left(\frac{\text{AFT}_{it}}{A_{it-1}}\right) + \epsilon_{it} \quad [9]$$

Nota: TAit - Acréscimos totais no ano t para a empresa i; ΔR_{it} - Receitas no ano t menos receitas no ano t-1 para a empresa i; AFT_{it} - Ativos fixos brutos, instalações e equipamentos no ano t para a empresa i; A_{it-1} - Ativos totais no ano t-1 para a empresa i; ϵ_{it} - Termo de erro no ano t para a empresa i. Fonte: Elaboração própria, com recurso a Costa e Soares, (2022) e Jones (1991).

Os coeficientes estimados na regressão geral, indicam de que forma as variáveis explicativas influenciam os AT normalizados das empresas. O coeficiente do valor constante ($\alpha_0 = 0,34497$) representa o valor médio esperado dos AT, em relação aos ativos do período anterior, quando as variáveis explicativas assumem um valor nulo. Este valor reflete um ponto de partida dos *accruals*, indicando que, mesmo na ausência de variações nas receitas, nos ativos fixos ou de efeitos de escala, existe um patamar mínimo esperado de *accruals* associados às operações da empresa. A presença de uma constante positiva sugere que as empresas, em média, apresentam *accruals* estruturais, decorrentes de práticas normais de reconhecimento contabilístico, independentemente das demais variáveis explicativas do modelo.

O coeficiente α_1 correspondente a -8825,32 capta diferenças de escala entre empresas, sugerindo que empresas com maior volume de ativos apresentam, em média, *accruals* normalizados inferiores.

O coeficiente β_1 referente ao valor de -0,11804, também é negativo, o que indica que os aumentos na receita estão relacionados com reduções nos *accruals*, o que sugere que a variação das vendas reflete sobretudo desempenho operacional real e não práticas de manipulação contabilística.

Por fim, β_2 correspondente a -0,20473, apresenta sinal negativo, o que indica que as empresas com maior proporção de AFT tendem a evidenciar níveis inferiores de AT, possivelmente em virtude do maior efeito das despesas de depreciação associadas a este tipo de ativos.

Deste modo, os valores estimados permitem isolar a componente não discricionária dos *accruals*, explicada pelas variações nas receitas e pelo nível de ativos, sendo o termo de erro (ϵ_{it}) interpretado como uma aproximação aos *accruals* discricionários, os quais poderão refletir eventuais práticas de manipulação de resultados.

Seguidamente, procede-se à estimação do modelo de *accruals* para cada ano individualmente, com o objetivo de captar possíveis variações temporais na relação entre os AT e as variáveis explicativas. Esta abordagem permite analisar se os determinantes dos *accruals* normais se mantêm estáveis ao longo do tempo, ou se apresentam alterações significativas associadas a diferentes contextos económicos ou contabilísticos. Assim, as regressões anuais possibilitam identificar eventuais mudanças no comportamento dos *accruals* não discricionários e na magnitude dos *accruals* discricionários, refletindo potenciais diferenças na prática de manipulação de resultados entre exercícios.

Tabela 8 - Regressão linear às variáveis em análise, por ano

ANO	α_0	α_1	β_1	β_2	P-Value α_0	Nível de significância	P-Value α_1	Nível de significância	P-Value β_1	Nível de significância	P-Value β_2	Nível de significância	R^2	R^2 adj	P-Value global	Nível de significância	n
2020	0,853	-42023,657	-0,887	0,874	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,6935	0,6934	0,000	<0,05	22420
2021	0,093	-13203,591	-0,043	0,851	0,169	>=0,05	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,3525	0,3524	0,000	<0,05	22714
2022	0,351	-5398,279	-0,074	-0,785	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,4347	0,4346	0,000	<0,05	22727
2023	0,831	-8406,397	-0,166	-1,467	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,000	<0,05	0,6905	0,6904	0,000	<0,05	21356

Fonte: Elaboração própria, com recurso aos softwares Rstudio e ao Excel.

A Tabela 8 apresenta os resultados da regressão linear aplicada às variáveis em análise para o período de 2020 a 2023. A análise das regressões anuais evidencia algumas variações relevantes nos coeficientes estimados ao longo do período em estudo. O valor do termo constante (α_0) apresenta oscilações entre os diferentes anos (0,853; 0,093; 0,351; 0,831), o que sugere alterações no nível médio de *accruals* não discricionários das empresas, possivelmente decorrentes de condições económicas distintas ou de mudanças nos pressupostos contabilísticos.

O coeficiente associado ao inverso dos ativos totais (α_1) surge com coeficientes muito grandes e negativos (-42023,657; -13203,591; -5398,279; -8406,397). Isto costuma indicar problemas de escala/unidade, por exemplo, ativos registados em valores nominais muito elevados enquanto as restantes variáveis já se encontram escaladas, ou heterogeneidade extrema de dimensão entre empresas.

O coeficiente relativo à variação das receitas (β_1) é sempre negativo, mas com uma magnitude reduzida em 2021 e 2022 e mais forte em 2020 e 2023 (-0,887 em 2020; -0,043 em 2021; -0,074 em 2022; -0,166 em 2023). Isto sugere que a variação da receita (ou do volume de negócios) tende a associar-se a reduções dos *accruals*, o que contraria o sinal mais comum no modelo de Jones, salvo em 2021 e

2022, anos em que o efeito é praticamente nulo. Esta evidência reforça a hipótese de efeitos da pandemia, subsídios e diferimentos nesses anos, ou de problemas de escala/medição das receitas (por exemplo, receita bruta vs. receita líquida, efeitos de inflação ou reclassificações).

Por um lado, o coeficiente dos ativos fixos tangíveis (β_2) muda ao longo do tempo, em 2020 (+0,874) e 2021 (+0,851) a relação é positiva, o que contraria o esperado no modelo de Jones, em que o coeficiente costuma ser negativo porque mais PPE implica mais depreciação, logo menores *accruals*. Por outro, em 2022 (-0,785) e 2023 (-1,467) o sinal volta ao padrão esperado, com magnitude especialmente forte em 2023. Ou seja, 2020 e 2021 podem refletir distorções da pandemia e/ou choques sectoriais (alterações contabilísticas, subsídios, diferimentos) que inverteram o padrão, ao passo que 2022 e 2023 sugerem normalização e um forte componente económico dos *accruals* ligado ao ativo fixo tangível.

Esta alteração sugere uma mudança no efeito dos ativos fixos sobre os AT: enquanto nos primeiros anos um maior peso relativo de ativos fixos esteve associado a níveis superiores de *accruals*, nos anos seguintes essa relação tornou-se negativa, possivelmente em virtude do impacto acrescido das despesas de depreciação no reconhecimento contabilístico.

Para cada coeficiente do modelo estimado, definiram-se as seguintes hipóteses:

- Para os coeficientes individualmente:

H0: O coeficiente é igual a zero. ($H_0: \alpha_0=0; H_0:\alpha_1=0; H_0:\beta_1=0; H_0:\beta_2=0$).

H1: O coeficiente não é igual a zero. ($H_0: \alpha_0 \neq 0; H_0: \alpha_1 \neq 0; H_0: \beta_1 \neq 0; H_0: \beta_2 \neq 0$).

- Para a significância conjunta do modelo:

H0: Todos os coeficientes das variáveis explicativas são iguais a zero. ($H_0: \alpha_0=\alpha_1=\beta_1=\beta_2=0$).

H1: Pelo menos um dos coeficientes é diferente de zero.

De forma geral, os *p-values* de α_1, β_1 e β_2 são inferiores a 5%, pelo que se rejeita H_0 e conclui-se que estes parâmetros são estatisticamente significativos em todos os anos. A exceção é o interceto α_0 em 2021, com *p-value* = 0,169 ($\geq 0,05$), para o qual não se rejeita H_0 .

O *p-value* global é 0,000 em todos os anos, sendo inferior a 5%, logo rejeita-se H_0 , concluindo que as variáveis independentes têm capacidade explicativa sobre a variável dependente, ou seja, têm poder explicativo conjunto.

Relativamente ao coeficiente de determinação (R^2) e ao R^2 ajustado, observam-se valores elevados em 2020 e 2023 (cerca de 0,69), moderados em 2022 (0,43) e mais baixos em 2021 (0,35). Estes resultados indicam que o modelo explica uma proporção considerável da variabilidade da variável dependente, sobretudo em 2020 e 2023, enquanto em 2021 e 2022 a capacidade explicativa é mais limitada. O facto

de o R^2 ajustado ser muito próximo do R^2 evidencia uma estrutura equilibrada do modelo, indicando que o número de variáveis explicativas é proporcional à dimensão da amostra.

Em termos globais, os resultados não evidenciam uma estabilidade plena, pelo contrário, revelam variações relevantes entre anos, consistentes com choques temporários em 2020 e 2021 e uma normalização progressiva em 2022 e 2023. Assim, a relação entre os *accruals* e as variáveis explicativas mostra-se sensível ao contexto económico e às políticas contabilísticas adotadas em cada período.

3.4. Análise dos Índícios de Manipulação de Resultados com Incentivo da Minimização do Pagamento de Impostos ao Estado

Após estimar os *accruals* discricionários para cada empresa e período, foi incluída a variável LBC, à base de dados, definida como *dummy* igual a 1 quando a observação apresenta combinações críticas na aplicação da Lei de Benford e 0 nos restantes casos. Esta inclusão, visa verificar se as entidades com desvios relevantes face à distribuição teórica de Benford, apresentam também níveis superiores de *accruals* discricionários absolutos (DACA), permitindo explorar uma possível ligação entre anomalias nos padrões numéricos e eventuais práticas de manipulação contabilística.

A inclusão desta variável possibilitou examinar se as entidades que evidenciaram desvios significativos à distribuição teórica da Lei de Benford, apresentavam também níveis superiores de *accruals* discricionários absolutos (DACA), o que poderia sugerir uma eventual associação entre anomalias nos padrões numéricos e práticas de manipulação contabilística.

O objetivo é avaliar indícios de manipulação de resultados relacionados com a minimização da carga fiscal. Para tal, procede-se à análise da correlação de Pearson e a correlação de Spearman, calculada anualmente e por quartis de dimensão, como verificação de robustez.

Tabela 9 - Correlação de Pearson e de Correlação de Spearman por ano

Correlação de Pearson									
Para 2020:					Para 2021:				
	LBC	DACA	NDIM	ISRI		LBC	DACA	NDIM	ISRI
LBC	1,0000	-0,0021	-0,0257	0,0051	LBC	1,0000	0,0035	-0,0341	-0,0052
P-Value	NA	0,7565	0,0001	0,4432	P-Value	NA	0,6005	0,0000	0,4336
N	22420	22420	22420	22420	N	22714	22714	22714	22714
DACA	-0,0021	1,0000	-0,0053	-0,0017	DACA	0,0035	1,0000	-0,0077	-0,0032
P-Value	0,7565	NA	0,4285	0,8023	P-Value	0,6005	NA	0,2454	0,6342
N	22420	22420	22420	22420	N	22714	22714	22714	22714
NDIM	-0,0257	-0,0053	1,0000	0,0733	NDIM	-0,0341	-0,0077	1,0000	0,2647
P-Value	0,0001	0,4285	NA	0,0000	P-Value	0,0000	0,2454	NA	0,0000
N	22420	22420	22420	22420	N	22714	22714	22714	22714
ISRI	0,0051	-0,0017	0,0733	1,0000	ISRI	-0,0052	-0,0032	0,2647	1,0000
P-Value	0,4432	0,8023	0,0000	NA	P-Value	0,4336	0,6342	0,0000	NA
N	22420	22420	22420	22420	N	22714	22714	22714	22714

	Para 2022:				Para 2023:			
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI
LBC	1,0000	-0,0055	-0,0472	-0,0028	LBC	1,0000	0,0022	-0,0392
P-Value	NA	0,4076	0,0000	0,6727	P-Value	NA	0,7494	0,8255
N	22727	22727	22727	22727	N	21356	21356	21356
DACA	-0,0055	1,0000	-0,0196	-0,0051	DACA	0,0022	1,0000	-0,0084
P-Value	0,4076	NA	0,0031	0,4417	P-Value	0,7494	NA	0,2202
N	22727	22727	22727	22727	N	21356	21356	21356
NDIM	-0,0472	-0,0196	1,0000	0,2080	NDIM	-0,0392	-0,0084	1,0000
P-Value	0,0000	0,0031	NA	0,0000	P-Value	0,0000	0,2202	NA
N	22727	22727	22727	22727	N	21356	21356	21356
ISRI	-0,0028	-0,0051	0,2080	1,0000	ISRI	-0,0015	-0,0017	0,2031
P-Value	0,6727	0,4417	0,0000	NA	P-Value	0,8255	0,8094	0,0000
N	22727	22727	22727	22727	N	21356	21356	21356

Correlação de Spearman

	Para 2020:				Para 2021:			
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI
LBC	1,0000	-0,0100	-0,0309	-0,0014	LBC	1,0000	-0,0154	-0,0383
P-Value	NA	0,1351	0,0000	0,8293	P-Value	NA	0,0200	0,0000
N	22420	22420	22420	22420	N	22714	22714	22714
DACA	-0,0100	1,0000	0,0141	-0,0630	DACA	-0,0154	1,0000	-0,0009
P-Value	0,1351	NA	0,0346	0,0000	P-Value	0,0200	NA	0,8941
N	22420	22420	22420	22420	N	22714	22714	22714
NDIM	-0,0309	0,0141	1,0000	0,1122	NDIM	-0,0383	-0,0009	1,0000
P-Value	0,0000	0,0346	NA	0,0000	P-Value	0,0000	0,8941	NA
N	22420	22420	22420	22420	N	22714	22714	22714
ISRI	-0,0014	-0,0630	0,1122	1,0000	ISRI	0,0061	-0,1521	0,1491
P-Value	0,8293	0,0000	0,0000	NA	P-Value	0,3578	0,0000	0,0000
N	22420	22420	22420	22420	N	22714	22714	22714

	Para 2022:				Para 2023:			
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI
LBC	1,0000	0,0076	-0,0541	-0,0179	LBC	1,0000	0,0074	-0,0472
P-Value	NA	0,2494	0,0000	0,0071	P-Value	NA	0,2819	0,1182
N	22727	22727	22727	22727	N	21356	21356	21356
DACA	0,0076	1,0000	-0,0716	-0,2027	DACA	0,0074	1,0000	-0,0585
P-Value	0,2494	NA	0,0000	0,0000	P-Value	0,2819	NA	0,0000
N	22727	22727	22727	22727	N	21356	21356	21356
NDIM	-0,0541	-0,0716	1,0000	0,2096	NDIM	-0,0472	-0,0585	1,0000
P-Value	0,0000	0,0000	NA	0,0000	P-Value	0,0000	0,0000	NA
N	22727	22727	22727	22727	N	21356	21356	21356
ISRI	-0,0179	-0,2027	0,2096	1,0000	ISRI	-0,0107	-0,1628	0,2450
P-Value	0,0071	0,0000	0,0000	NA	P-Value	0,1182	0,0000	0,0000
N	22727	22727	22727	22727	N	21356	21356	21356

Nota: LBC – Lei de Benford crítico; DACA – *Accruals* discricionários absolutos; NDIM – Dimensão das empresas; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

De acordo com a Tabela 9, os resultados indicam, de forma consistente, ausência de relação material entre a LBC e os DACA: os coeficientes são próximos de zero em todos os anos, nas duas correlações, e quando aparecem estatisticamente significativos tal decorre, muito provavelmente, do tamanho amostra elevado, não tendo relevância prática.

Entre DACA e NDIM (dimensão), as correlações são igualmente muito reduzidas e sem padrão estável, os sinais alternam e magnitudes $< |0,02|$ em Pearson, e em Spearman surgem valores modestos, negativos em 2022 e 2023, o que sugere fraca associação entre a variabilidade dos *accruals* discricionários e a dimensão empresarial.

Entre DACA e ISRI, as correlações de Pearson são praticamente nulas em todos os anos e em Spearman observam-se correlações negativas e pequenas (mais visíveis em 2021 e 2023), sinalizando uma associação monotônica fraca e sem implicações claras.

A única relação que se destaca é a de NDIM com ISRI, que surge positiva e estatisticamente significativa em todos os anos, moderada em 2020 ($\approx 0,07$) na correlação de Pearson e em 2021 e 2023 ($\approx 0,15$ e $0,25$ em Spearman; $\approx 0,26$ e $0,20$ em Pearson), é compatível com maiores impostos pagos por empresas de maior dimensão.

Em síntese, não se identifica evidência de associação sistemática entre LBC e DACA, e as restantes correlações, quando significativas, apresentam magnitudes baixas, devendo ser interpretadas com cautela.

Face à ausência de evidência conclusiva na Tabela 9, procedeu-se à estratificação anual dos DAC em quartis. A Tabela 10, com o primeiro quartil (Q1) a reunir os valores mais negativos e o quartil quatro (Q4) os mais positivos, para observar com maior detalhe a incidência da LBC. No total de 2020 a 2023, a proporção de observações com LBC=1 é de cerca de 17,1% (15 232 em 89 217 dados), distribuindo-se por ano em 3 808, 3 819, 3 988 e 3 617 casos, respetivamente. Os intervalos dos DAC por quartil confirmam a ordenação esperada (por exemplo: 2020: Q1 -2167,8 a -1,73; Q4 -0,18 a 1164,5; 2023: Q1 -1377,5 a -0,21; Q4 -0,26 a 1984,7), evidenciando extremos mais acentuados em 2020 e 2023, encontra-se apresentados no Apêndice 5.

Tabela 10 - Correlação de Pearson por ano e por quartil

Correlação de Pearson																	
Ano	Para Q1:				Para Q2:				Para Q3:				Para Q4:				
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	
2020	LBC	1,0000	-0,0083	-0,0315	0,0219	1,0000	-0,0085	-0,0462	-0,0056	1,0000	0,0077	-0,0050	0,0159	1,0000	0,0046	-0,0023	0,0053
	P-Value	NA	0,5328	0,0182	0,1016	NA	0,5251	0,0005	0,6728	NA	0,5630	0,7076	0,2332	NA	0,7307	0,8621	0,6925
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
	DACA	-0,0083	1,0000	-0,0045	0,0008	-0,0085	1,0000	0,0223	0,0152	0,0077	1,0000	0,0495	0,0406	0,0046	1,0000	-0,0041	-0,0081
	P-Value	0,5328	NA	0,7335	0,9552	0,5251	NA	0,0944	0,2548	0,5630	NA	0,0002	0,0024	0,7307	NA	0,7596	0,5459
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
	NDIM	-0,0315	-0,0045	1,0000	-0,3251	-0,0462	0,0223	1,0000	0,2660	-0,0050	0,0495	1,0000	0,5963	-0,0023	-0,0041	1,0000	0,4139
	P-Value	0,0182	0,7335	NA	0,0000	0,0005	0,0944	NA	0,0000	0,7076	0,0002	NA	0,0000	0,8621	0,7596	NA	0,0000
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
	ISRI	0,0219	0,0008	-0,3251	1,0000	-0,0056	0,0152	0,2660	1,0000	0,0159	0,0406	0,5963	1,0000	0,0053	-0,0081	0,4139	1,0000
	P-Value	0,1016	0,9552	0,0000	NA	0,6728	0,2548	0,0000	NA	0,2332	0,0024	0,0000	NA	0,6925	0,5459	0,0000	NA
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
2021	LBC	1,0000	-0,0029	-0,0439	-0,0065	1,0000	0,0075	-0,0470	-0,0099	1,0000	-0,0040	-0,0197	-0,0072	1,0000	0,0236	-0,0166	0,0090
	P-Value	NA	0,8267	0,0009	0,6258	NA	0,5701	0,0004	0,4575	NA	0,7618	0,1368	0,5893	NA	0,0760	0,2101	0,4989
	N	5679	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678
	DACA	-0,0029	1,0000	-0,0078	-0,0023	0,0075	1,0000	0,0255	-0,0120	-0,0040	1,0000	0,0165	0,0228	0,0236	1,0000	-0,0080	-0,0277
	P-Value	0,8267	NA	0,5588	0,8621	0,5701	NA	0,0543	0,3652	0,7618	NA	0,2133	0,0855	0,0760	NA	0,5489	0,0369
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
	NDIM	-0,0439	-0,0078	1,0000	0,3202	-0,0470	0,0255	1,0000	0,3431	-0,0197	0,0165	1,0000	0,2473	-0,0166	-0,0080	1,0000	0,1889
	P-Value	0,0009	0,5588	NA	0,0000	0,0004	0,0543	NA	0,0000	0,1368	0,2133	NA	0,0000	0,2101	0,5489	NA	0,0000
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
	ISRI	-0,0065	-0,0023	0,3202	1,0000	-0,0099	-0,0120	0,3431	1,0000	-0,0072	0,0228	0,2473	1,0000	0,0090	-0,0277	0,1889	1,0000
	P-Value	0,6258	0,8621	0,0000	NA	0,4575	0,3652	0,0000	NA	0,5893	0,0855	0,0000	NA	0,4989	0,0369	0,0000	NA
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
2022	LBC	1,0000	-0,0039	-0,0355	0,0059	1,0000	-0,0017	-0,0588	-0,0068	1,0000	0,0184	-0,0566	-0,0045	1,0000	-0,0144	-0,0335	-0,0099
	P-Value	NA	0,7686	0,0075	0,6558	NA	0,8979	0,0000	0,6066	NA	0,1649	0,0000	0,7366	NA	0,2782	0,0115	0,4566
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681

Correlação de Pearson																	
Ano	Para Q1:					Para Q2:				Para Q3:				Para Q4:			
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	
2022	DACA	-0,0039	1,0000	-0,0228	-0,0128	-0,0017	1,0000	0,0162	0,0212	0,0184	1,0000	-0,0133	0,0123	-0,0144	1,0000	-0,0125	-0,0099
	P-Value	0,7686	NA	0,0856	0,3341	0,8979	NA	0,2212	0,1103	0,1649	NA	0,3170	0,3524	0,2782	NA	0,3475	0,4570
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
	NDIM	-0,0355	-0,0228	1,0000	0,3475	-0,0588	0,0162	1,0000	0,4624	-0,0566	-0,0133	1,0000	0,1390	-0,0335	-0,0125	1,0000	0,1501
	P-Value	0,0075	0,0856	NA	0,0000	0,0000	0,2212	NA	0,0000	0,0000	0,3170	NA	0,0000	0,0115	0,3475	NA	0,0000
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
2022	ISRI	0,0059	-0,0128	0,3475	1,0000	-0,0068	0,0212	0,4624	1,0000	-0,0045	0,0123	0,1390	1,0000	-0,0099	-0,0099	0,1501	1,0000
	P-Value	0,6558	0,3341	0,0000	NA	0,6066	0,1103	0,0000	NA	0,7366	0,3524	0,0000	NA	0,4566	0,4570	0,0000	NA
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
	LBC	1,0000	0,0044	-0,0288	0,0028	1,0000	-0,0061	-0,0470	-0,0019	1,0000	0,0306	-0,0359	0,0082	1,0000	0,0017	-0,0491	-0,0040
	P-Value	NA	0,7453	0,0352	0,8369	NA	0,6569	0,0006	0,8903	NA	0,0252	0,0087	0,5485	NA	0,9021	0,0003	0,7683
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339
2023	DACA	0,0044	1,0000	-0,0096	-0,0062	-0,0061	1,0000	0,0019	0,0129	0,0306	1,0000	-0,0061	-0,0029	0,0017	1,0000	-0,0076	-0,0011
	P-Value	0,7453	NA	0,4812	0,6510	0,6569	NA	0,8869	0,3474	0,0252	NA	0,6540	0,8349	0,9021	NA	0,5765	0,9341
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339
	NDIM	-0,0288	-0,0096	1,0000	0,3632	-0,0470	0,0019	1,0000	0,3624	-0,0359	-0,0061	1,0000	0,1625	-0,0491	-0,0076	1,0000	0,2041
	P-Value	0,0352	0,4812	NA	0,0000	0,0006	0,8869	NA	0,0000	0,0087	0,6540	NA	0,0000	0,0003	0,5765	NA	0,0000
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339
2023	ISRI	0,0028	-0,0062	0,3632	1,0000	-0,0019	0,0129	0,3624	1,0000	0,0082	-0,0029	0,1625	1,0000	-0,0040	-0,0011	0,2041	1,0000
	P-Value	0,8369	0,6510	0,0000	NA	0,8903	0,3474	0,0000	NA	0,5485	0,8349	0,0000	NA	0,7683	0,9341	0,0000	NA
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339

Correlação de Spearman																	
Ano	Para Q1:					Para Q2:				Para Q3:				Para Q4:			
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	
2020	LBC	1,0000	-0,0135	-0,0378	0,0000	1,0000	-0,0082	-0,0481	-0,0077	1,0000	0,0058	-0,0175	0,0035	1,0000	-0,0120	-0,0023	0,0002
	P-Value	NA	0,3131	0,0047	0,9982	NA	0,5385	0,0003	0,5621	NA	0,6656	0,1894	0,7907	NA	0,3687	0,8621	0,9894
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
	DACA	-0,0135	1,0000	-0,0444	-0,1412	-0,0082	1,0000	0,0182	-0,0151	0,0058	1,0000	0,0565	0,0812	-0,0120	1,0000	-0,0268	-0,0984
	P-Value	0,3131	NA	0,0009	0,0000	0,5385	NA	0,1736	0,2573	0,6656	NA	0,0000	0,0000	0,3687	NA	0,0444	0,0000
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
2020	NDIM	-0,0378	-0,0444	1,0000	0,0755	-0,0481	0,0182	1,0000	0,1484	-0,0175	0,0565	1,0000	0,1438	-0,0023	-0,0268	1,0000	0,0663
	P-Value	0,0047	0,0009	NA	0,0000	0,0003	0,1736	NA	0,0000	0,1894	0,0000	NA	0,0000	0,8621	0,0444	NA	0,0000
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
	ISRI	0,0000	-0,1412	0,0755	1,0000	-0,0077	-0,0151	0,1484	1,0000	0,0035	0,0812	0,1438	1,0000	0,0002	-0,0984	0,0663	1,0000
	P-Value	0,9982	0,0000	0,0000	NA	0,5621	0,2573	0,0000	NA	0,7907	0,0000	0,0000	NA	0,9894	0,0000	0,0000	NA
	N	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605	5605
2021	LBC	1,0000	-0,0004	-0,0485	0,0061	1,0000	0,0086	-0,0553	-0,0016	1,0000	-0,0047	-0,0179	-0,0082	1,0000	-0,0210	-0,0166	0,0275
	P-Value	NA	0,9782	0,0003	0,6452	NA	0,5150	0,0000	0,9037	NA	0,7205	0,1782	0,5347	NA	0,1131	0,2101	0,0386
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
	DACA	-0,0004	1,0000	-0,0475	-0,1755	0,0086	1,0000	0,0225	0,0023	-0,0047	1,0000	0,0197	0,0426	-0,0210	1,0000	-0,0382	-0,1563
	P-Value	0,9782	NA	0,0003	0,0000	0,5150	NA	0,0894	0,8603	0,7205	NA	0,1386	0,0013	0,1131	NA	0,0040	0,0000
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
2021	NDIM	-0,0485	-0,0475	1,0000	0,1446	-0,0553	0,0225	1,0000	0,2240	-0,0179	0,0197	1,0000	0,1016	-0,0166	-0,0382	1,0000	0,0543
	P-Value	0,0003	0,0003	NA	0,0000	0,0000	0,0894	NA	0,0000	0,1782	0,1386	NA	0,0000	0,2101	0,0040	NA	0,0000
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
	ISRI	0,0061	-0,1755	0,1446	1,0000	-0,0016	0,0023	0,2240	1,0000	-0,0082	0,0426	0,1016	1,0000	0,0275	-0,1563	0,0543	1,0000
	P-Value	0,6452	0,0000	0,0000	NA	0,9037	0,8603	0,0000	NA	0,5347	0,0013	0,0000	NA	0,0386	0,0000	0,0000	NA
	N	5679	5679	5679	5679	5679	5679	5679	5679	5678	5678	5678	5678	5678	5678	5678	5678
2022	LBC	1,0000	0,0157	-0,0455	-0,0081	1,0000	-0,0017	-0,0659	-0,0120	1,0000	0,0171	-0,0646	-0,0079	1,0000	0,0121	-0,0338	-0,0429
	P-Value	NA	0,2376	0,0006	0,5438	NA	0,8994	0,0000	0,3660	NA	0,1967	0,0000	0,5506	NA	0,3611	0,0061	0,0012
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
	DACA	0,0157	1,0000	-0,1199	-0,3984	-0,0017	1,0000	0,0159	0,0555	0,0171	1,0000	-0,0170	-0,0509	0,0121	1,0000	-0,0364	-0,1130
	P-Value	0,2376	NA	0,0000	0,0000	0,8994	NA	0,2320	0,0000	0,1967	NA	0,1996	0,0001	0,3611	NA	0,0061	0,0000
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
2022	NDIM	-0,0455	-0,1199	1,0000	0,2216	-0,0659	0,0159	1,0000	0,2444	-0,0646	-0,0170	1,0000	0,2120	-0,0338	-0,0364	1,0000	0,1281
	P-Value	0,0006	0,0000	NA	0,0000	0,0000	0,2320	NA	0,0000	0,0000	0,1996	NA	0,0000	0,0107	0,0061	NA	0,0000
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
	ISRI	-0,0081	-0,3984	0,2216	1,0000	-0,0120	0,0555	0,2444	1,0000	-0,0079	-0,0509	0,2120	1,0000	-0,0429	-0,1130	0,1281	1,0000
	P-Value	0,5438	0,0000	0,0000	NA	0,3660	0,0000	0,0000	NA	0,5506	0,0001	0,0000	NA	0,0012	0,0000	0,0000	NA
	N	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5682	5681	5681	5681	5681
2023	LBC	1,0000	0,0260	-0,0379	-0,0190	1,0000	-0,0054	-0,0559	-0,0265	1,0000	0,0306	-0,0466	0,0062	1,0000	-0,0028	-0,0515	-0,0003
	P-Value	NA	0,0575	0,0056	0,1653	NA	0,6958	0,0000	0,0529	NA	0,0252	0,0007	0,6519	NA	0,8371	0,0002	0,9805
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339
	DACA	0,0260	1,0000	-0,0935	-0,3861	-0,0054	1,0000	0,0018	0,0254	0,0306	1,0000	-0,0066	0,0163	-0,0028	1,0000	-0,0853	-0,1342
	P-Value	0,0575	NA	0,0000	0,0000	0,6958	NA	0,8957	0,0631	0,0252	NA	0,6318	0,2351	0,8371	NA	0,0000	0,0000
	N	5339															

Correlação de Spearman																	
	Para Q1:					Para Q2:					Para Q3:				Para Q4:		
	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	LBC	DACA	NDIM	ISRI	
2023	NDIM	-0,0379	-0,0935	1,0000	0,2174	-0,0559	0,0018	1,0000	0,3058	-0,0466	-0,0066	1,0000	0,2552	-0,0515	-0,0853	1,0000	0,1662
	P.Value	0,0056	0,0000	NA	0,0000	0,0000	0,8957	NA	0,0000	0,0007	0,6318	NA	0,0000	0,0002	0,0000	NA	0,0000
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339
	ISRI	-0,0190	-0,3861	0,2174	1,0000	-0,0265	0,0254	0,3058	1,0000	0,0062	0,0163	0,2552	1,0000	-0,0003	-0,1342	0,1662	1,0000
	P.Value	0,1653	0,0000	0,0000	NA	0,0529	0,0631	0,0000	NA	0,6519	0,2351	0,0000	NA	0,9805	0,0000	0,0000	NA
	N	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339	5339

Nota: LBC – Lei de Benford crítico; DACA – *Accruals* discricionários absolutos; NDIM – Dimensão das empresas; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

Com base nas matrizes de correlação de Pearson e de Spearman, Tabela 10, calculadas anualmente e por quartis em função da variável DAC (sendo DACA o valor absoluto dos *accruals* discricionários), a leitura global é clara e consistente. Entre 2020 e 2023 observa-se de forma predominante uma associação positiva entre NDIM e ISRI, fraca a moderada em Pearson e monotónica em Spearman, coerente com um efeito de escala em que empresas maiores apresentam maiores montantes de imposto e sem constituir indício por si só de manipulação de resultados.

Em conjunto, a leitura cruzada das correlações de Pearson e Spearman reforça que a relação entre DACA, NDIM e ISRI não é uniforme, variando em intensidade e sinal conforme o quartil e o tipo de correlação adotado. De modo geral, as matrizes mostram entre DACA com NDIM e DACA com ISRI relações fracas, isto é, as variáveis não se movem fortemente em conjunto. No seu todo, quando todas as correlações são positivas (> 0), implica uma tendência direta (à medida que DACA aumenta, NDIM e ISRI tendem a aumentar), conceitualmente, correlações negativas traduziram uma tendência inversa (DACA sobe enquanto NDIM e ISRI descem). Nas subamostras 2020 Q3, 2021 Q1, 2021 Q4 e 2022 Q1, para correlação de Pearson e nas subamostras 2020 Q1, Q3 e Q4, 2021 Q1, Q3 e Q4, 2022 para todos os quartis e 2023 Q1 e Q4, na correlação de Spearman, na qual se verifica que as correlações foram estatisticamente significativas ($p\text{-value} < 0,05$), indicando que, nesses períodos, a associação entre DACA e as demais variáveis foi suficientemente consistente para rejeitar a hipótese de ausência de relação. As correlações positivas e significativas entre DACA, NDIM e ISRI indicam que, nesses períodos, DACA aumenta em conjunto com esses indicadores, já correlações negativas e significativas entre DACA e ISRI são compatíveis com cenários de minimização fiscal via gestão de resultados (maior *accruals* associado a um menor imposto do período). Assim, ainda que não se observe um padrão temporal contínuo, a persistência de sinais significativos em certos quartis sustenta a relevância analítica das correlações, apontando para episódios de possível gestão fiscal associada a *accruals* e para variação estrutural na relação entre desempenho contabilístico e fiscal.

No que respeita ao LBC, as correlações com ISRI, NDIM e DACA são, no agregado, muito próximas de zero em ambos os coeficientes. Isto significa que as anomalias de Benford não se mostram associadas, de forma linear ou monotónica, às referidas variáveis quando se observa o universo de empresas e anos. A exceção mais relevante surge em 2023, quartil 3, quando se verifica sinal positivo

entre LBC e DACA ($r \approx 0,031$) e sinal negativo entre LBC e NDIM ($r \approx 0,036$), ambos estatisticamente significativos ($p < 0,05$). Este padrão direcional oposto destaca-se como um marco analítico, o que indica associações consistentes e diferenciadas de LBC com DACA e com NDIM, sugerindo dinâmicas específicas nesse período. Em síntese, 2023 Q3 revela um padrão estatisticamente significativo, no qual o LBC apresenta relação positiva com DACA e negativa com NDIM, evidenciando divergência apenas no comportamento face ao NDIM, o que confere relevância analítica e interpretativa ao período em análise. Este resultado sugere a possibilidade de manipulação de resultados, uma vez que o aumento dos *accruals* discricionários (DACA) coincide com maior desvio à Lei de Benford, indicando anomalias numéricas associadas a práticas contabilísticas

Em síntese, a evidência combinada de Pearson e Spearman não aponta para um padrão generalizado de manipulação de resultados com vista à minimização da carga fiscal quando se observa o conjunto de anos e quartis. Todavia, existem indícios consistentes e economicamente plausíveis de tal comportamento nos quartis extremos (Q1 e Q4) ao longo de vários anos, onde a associação negativa dos DACA com ISRI é repetida e, por vezes, mais intensa. Fora desses casos, e em particular nos quartis intermédios (Q2 e Q3), as correlações permanecem fracas. Globalmente, quaisquer evidências do fenómeno parecem esporádicas e delimitadas a segmentos específicos. Importa, por isso, aprofundar a análise nos anos/quartis em que a relação os DACA com ISRI se revela negativa e significativa. Nos restantes, os resultados alinham-se com um funcionamento regular e com efeitos de escala.

Neste contexto, a combinação dos resultados obtidos através do modelo de Jones e da Lei de Benford oferece uma visão complementar sobre o comportamento contabilístico das empresas do setor da restauração e similares. A possibilidade de evidência de manipulação dos resultados, associada à presença de alguns desvios localizados, sugere que as variações identificadas decorrem sobretudo de particularidades estruturais e operacionais do setor, e não de práticas generalizadas de distorção intencional da informação financeira.

3.5. Comparação e Integração dos Resultados

De forma a complementar, a leitura descritiva fornecida pelas correlações de Pearson e Spearman, recorre-se a um modelo *logit*, estimado no conjunto total de observações e separadamente por ano e por quadril, com o objetivo de avaliar, se as variáveis consideradas se associam à probabilidade de ocorrência de anomalia na LBC. A variável dependente é um indicador binário de LBC crítico (1 se a observação é sinalizada pela Lei de Benford crítico e 0 restantes casos). As variáveis explicativas são: DACA ($|DAC|$, magnitude absoluta dos *accruals* discricionários), NDIM (dimensão da empresa) e ISRI (imposto do período).

Esta abordagem é útil por três razões. Primeiro, permite verificar se as associações observadas nas correlações persistem quando controladas pelas restantes. Segundo, transforma a análise numa probabilidade diretamente interpretável, isto é, em que medida variações em DACA, NDIM ou ISRI se traduzem em maior ou menor risco de uma observação ser classificada como crítica pela Lei de Benford. Terceiro, ao estimar o modelo por ano, capta-se a possível heterogeneidade temporal do fenómeno.

Formalmente, o modelo assume:

$$P(\text{LBC} = 1 | X) = \text{logit}^{-1}(\beta_0 + \beta_1 \text{DACA} + \beta_2 \text{NDIM} + \beta_3 \text{ISRI}) \quad [14]$$

Nota: P – Probabilidade; X – Vetor de preditores; β_0 - Constante; $\beta_1, \beta_2, \beta_3$ - Coeficientes das covariáveis; LBC – Lei de Benford Crítico; DACA – *Accruals* discricionários absolutos; NDIM – Dimensão das empresas; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria.

No modelo *logit*, cada coeficiente, indica como uma variação na variável explicativa, se associa à probabilidade de ocorrência de LBC, mantendo as restantes variáveis constantes. Para facilitar a leitura, reportam-se *odds ratios* (razão de probabilidades) e/ou efeitos marginais (variação pontual/percentual na probabilidade), com intervalos de confiança (Szumilas, 2010). Sempre que relevante, discute-se a direção esperada dos sinais à luz da evidência anterior (por exemplo, se |DAC| tende ou não a aumentar a probabilidade de anomalia), bem como a consistência entre anos.

As correlações servem para detetar associações lineares (Pearson) e monotónicas robustas a *outliers* (Spearman), enquanto o modelo *logit* verifica, em quadro multivariado, se essas relações se traduzem numa maior ou menor probabilidade de ocorrência de anomalia na LBC, controlando pelas restantes variáveis. Assim, o *logit* complementa as correlações e permite avaliar, de forma integrada, se as *proxies* contabilísticas e fiscais estão associadas à LBC no universo estudado.

Com base no modelo *logit* estimado para cada ano, sabendo que a variável dependente é LBC, os resultados são consistentes no período de 2020 a 2023. A variável DACA apresenta coeficientes praticamente nulos e não significativos em todos os anos, indicando que, mantendo constantes a NDIM e o ISRI, as variações nos DACA não alteram a probabilidade de ocorrência de desvios relativamente à Lei de Benford.

Em todas as estimações por ano, a NDIM apresenta um coeficiente negativo e estatisticamente significativo ($p < 0,001$), com razões de probabilidades na ordem de 0,31 a 0,49 (2020: $e^{-0,950} \approx 0,39$; 2021: $e^{-1,165} \approx 0,31$; 2022: $e^{-1,108} \approx 0,33$; 2023: $e^{-0,721} \approx 0,49$). Isto implica uma redução dos *odds* de ser sinalizada como crítica entre aproximadamente 51% e 69% (2020: $1-0,39=61\%$) por cada aumento de uma unidade na NDIM, mantendo as restantes variáveis constantes. O efeito é consistente no tempo

e é compatível com a ideia de que empresas de maior dimensão tendem a dispor de processos e controlos mais robustos.

Já o ISRI apresenta coeficientes muito próximos de zero e não significativos, sugerindo que o montante do imposto não explica a LBC quando consideradas as restantes variáveis.

A constante é negativa e estatisticamente significativa em 2020, 2022 e 2023 ($p < 0,05$), refletindo uma probabilidade base do evento inferior a 50% quando as restantes variáveis valem zero (2020: $\frac{1}{1+e^{0,629}} \approx 0,35$; 2021: $\approx 0,40$; 2022: $\approx 0,40$; 2023: $\approx 0,30$). Esta variável, não altera a interpretação substantiva dos restantes coeficientes.

Em suma, o modelo confirma a leitura das correlações, ou seja, não existe evidência de que DACA ou ISRI estejam associados, de forma robusta, à ocorrência de LBC. Em contraste, a NDIM exerce influência clara, pois apresenta um coeficiente negativo e a variável é estatisticamente significativa, reduzindo a probabilidade de LBC. Conclui-se que não há indícios de um padrão sistemático de manipulação segundo a Lei de Benford.

Tabela 11 - Modelo *Logit* por ano

Logit por ano:						
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
2020	(Intercept)	1,000	-0,629	0,251	6,258	0,012
	DACA	1,000	0,000	0,001	0,107	0,744
	NDIM	1,000	-0,950	0,250	14,473	0,000
	ISRI	1,000	0,000	0,000	3,178	0,075
2021	(Intercept)	1,000	-0,425	0,243	3,065	0,080
	DACA	1,000	0,001	0,002	0,223	0,637
	NDIM	1,000	-1,165	0,241	23,399	0,000
	ISRI	1,000	0,000	0,000	0,265	0,606
2022	(Intercept)	1,000	-0,416	0,164	6,418	0,011
	DACA	1,000	-0,004	0,004	0,918	0,338
	NDIM	1,000	-1,108	0,161	47,080	0,000
	ISRI	1,000	0,000	0,000	2,892	0,089
2023	(Intercept)	1,000	-0,850	0,131	42,427	0,000
	DACA	1,000	0,000	0,001	0,071	0,790
	NDIM	1,000	-0,721	0,127	32,325	0,000
	ISRI	1,000	0,000	0,000	1,624	0,203

Nota: LBC – Lei de Benford crítico; DACA – *Accruals* discricionários absolutos; NDIM – Dimensão das empresas; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

A avaliação conjunta das correlações de Pearson e Spearman (por ano e por quartis de DAC) e dos modelos *logit* estimados por ano e por quartil permite traçar um retrato consistente sobre eventuais indícios de manipulação de resultados com vista à minimização da carga fiscal. As correlações servem para identificar associações lineares (Pearson) e monotónicas robustas a *outliers* (Spearman), enquanto

o *logit* verifica, num enquadramento multivariado, se essas relações se traduzem numa maior probabilidade de ocorrência de LBC, controlando por DACA (|DAC|), ISRI e NDIM.

Tabela 12 - Modelo *Logit* por ano e por quartil

Para 2020:							
	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n	
Q1	(Intercept)	1	-0,421	0,459	0,843	0,359	5605
	DACA	1	-0,009	0,012	0,531	0,466	5605
	ISRI	1	-1,142	0,456	6,282	0,012	5605
	NDIM	1	0,000	0,000	5,971	0,015	5605
Q2	(Intercept)	1	0,395	0,668	0,351	0,554	5605
	DACA	1	-0,123	0,218	0,320	0,572	5605
	ISRI	1	-1,808	0,590	9,403	0,002	5605
	NDIM	1	0,000	0,000	0,096	0,756	5605
Q3	(Intercept)	1	-0,960	0,516	3,454	0,063	5605
	DACA	1	0,075	0,130	0,333	0,564	5605
	ISRI	1	-0,646	0,508	1,615	0,204	5605
	NDIM	1	0,000	0,000	2,690	0,101	5605
Q4	(Intercept)	1	-1,169	1,211	0,931	0,335	5605
	DACA	1	0,000	0,001	0,120	0,729	5605
	ISRI	1	-0,441	1,212	0,132	0,716	5605
	NDIM	1	0,000	0,000	0,274	0,601	5605
Para 2021:							
	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n	
Q1	(Intercept)	1	-0,050	0,517	0,009	0,923	5679
	DACA	1	-0,001	0,003	0,058	0,810	5679
	ISRI	1	-1,550	0,514	9,087	0,003	5679
	NDIM	1	0,000	0,000	0,438	0,508	5679
Q2	(Intercept)	1	-0,566	0,364	2,416	0,120	5679
	DACA	1	0,225	0,339	0,439	0,508	5679
	ISRI	1	-1,124	0,336	11,155	0,001	5679
	NDIM	1	0,000	0,000	0,274	0,601	5679
Q3	(Intercept)	1	-0,834	0,487	2,939	0,086	5678
	DACA	1	-0,150	0,567	0,070	0,792	5678
	ISRI	1	-0,678	0,480	1,992	0,158	5678
	NDIM	1	0,000	0,000	0,172	0,678	5678
Q4	(Intercept)	1	10,507	186,852	0,003	0,955	5678
	DACA	1	0,005	0,003	2,824	0,093	5678
	ISRI	1	-12,155	186,852	0,004	0,948	5678
	NDIM	1	0,000	0,000	1,068	0,301	5678
Para 2022:							
	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n	
Q1	(Intercept)	1	-0,607	0,323	3,527	0,060	5682
	DACA	1	-0,002	0,005	0,118	0,731	5682
	ISRI	1	-0,950	0,319	8,885	0,003	5682
	NDIM	1	0,000	0,000	4,105	0,043	5682
Q2	(Intercept)	1	-0,219	0,306	0,512	0,474	5682
	DACA	1	-0,044	0,506	0,008	0,931	5682

		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q2	ISRI	1	-1,294	0,291	19,789	0,000	5682
	NDIM	1	0,000	0,000	5,536	0,019	5682
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q3	(Intercept)	1	-0,360	0,319	1,273	0,259	5682
	DACA	1	0,545	0,413	1,739	0,187	5682
	ISRI	1	-1,232	0,310	15,825	0,000	5682
	NDIM	1	0,000	0,000	0,063	0,802	5682
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q4	(Intercept)	1	0,087	0,701	0,016	0,901	5681
	DACA	1	-0,009	0,008	1,238	0,266	5681
	ISRI	1	-1,575	0,698	5,098	0,024	5681
	NDIM	1	0,000	0,000	0,446	0,504	5681
Para 2023:							
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q1	(Intercept)	1	-0,842	0,304	7,661	0,006	5339
	DACA	1	0,000	0,001	0,094	0,760	5339
	ISRI	1	-0,706	0,300	5,552	0,018	5339
	NDIM	1	0,000	0,000	2,173	0,140	5339
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q2	(Intercept)	1	-0,747	0,264	7,981	0,005	5339
	DACA	1	-0,155	0,343	0,205	0,650	5339
	ISRI	1	-0,748	0,210	12,638	0,000	5339
	NDIM	1	0,000	0,000	2,219	0,136	5339
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	n
Q3	(Intercept)	1	-1,011	0,265	14,505	0,000	5339
	DACA	1	1,133	0,514	4,871	0,027	5339
	ISRI	1	-0,731	0,251	8,463	0,004	5339
	NDIM	1	0,000	0,000	3,081	0,079	5339
		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	N
Q4	(Intercept)	1	1,191	1,030	1,337	0,248	5339
	DACA	1	0,000	0,001	0,009	0,926	5339
	ISRI	1	-2,782	1,029	7,312	0,007	5339
	NDIM	1	0,000	0,000	0,503	0,478	5339

Nota: LBC – Lei de Benford crítico; DACA – *Accruals* discricionários absolutos; NDIM – Dimensão das empresas; ISRI – Imposto sobre o rendimento do período. Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

Com base no modelo *logit* estimado por ano e por quartil de DAC, a variável DACA apresenta coeficientes muito próximos de zero e *p-value* elevados, o que indica ausência de efeito adicional nos *odds* de LBC quando comparamos entidades dentro do mesmo quartil.

Em contraste, a variável ISRI surge frequentemente com coeficientes negativos e estatisticamente significativos, traduzindo reduções relevantes dos *odds* de LBC por cada aumento de uma unidade em ISRI, ou seja, valores de imposto mais elevados, estão associados a uma menor probabilidade de sinalização crítica, mantendo as restantes variáveis constantes.

A variável NDIM, nos modelos por quartil, não é estimável de forma informativa, surgindo com estimativa e erro-padrão nulos, o que sugere falta de variabilidade ou colinearidade dentro de algumas células de ano-quartil e impede concluir sobre o seu efeito neste desenho. A constante é por vezes negativa e significativa, refletindo probabilidades base inferiores a 50% quando os preditores assumem valor zero, sem alterar a interpretação dos restantes coeficientes.

A variável NDIM representa a classe de dimensão (1 – microempresas; 2 – pequenas empresas; 3 – médias empresas; 4 – grandes empresas). No modelo por quartil, a variável NDIM surge sem efeito estimado (Estimate=0; SE=0), o que decorre de falta de variabilidade e/ou colinearidade dentro de vários quartis. Assim, não é possível inferir o efeito de NDIM nesta estratificação. Nos modelos anuais, onde há variabilidade suficiente, NDIM apresenta efeito protetor consistente.

No modelo *logit*, quartil 3 de 2023, verifica-se um coeficiente negativo e significativo ($\beta_0 = -1,011$; $p < 0,001$), o que corresponde a uma probabilidade base $\approx 27\%$ ($\frac{1}{1+e^{-(-1,011)}}$) de LBC=1 quando as restantes variáveis assumem o valor zero.

A variável DACA apresenta efeito positivo e estatisticamente significativo ($\beta_1 = 1,133$; $p = 0,027$), o *odds ratio* é $\approx 3,10$, ou seja, por cada aumento de uma unidade em DACA os *odds* de LBC aproximadamente triplicam, mantendo o resto das variáveis constantes.

A ISRI surge com efeito negativo e significativo ($\beta_3 = -0,731$; $p = 0,004$), implicando *odds ratio* é $\approx 0,48$ e, portanto, uma redução de 52% dos *odds* de LBC por unidade adicional.

A NDIM apresenta $\beta_2=0,000$ e $p=0,079$; ao nível padrão de 5% o efeito não é estatisticamente significativo. O padrão é compatível com baixa variabilidade ou colinearidade das classes de dimensão nesta amostra, pelo que o efeito é inconclusivo.

No Q3 de 2023, verifica-se associação positiva e estatisticamente significativa entre DACA e a probabilidade de LBC crítico ($OR \approx 3,10=e^{1,133}$), e associação negativa e significativa entre ISRI e essa probabilidade ($OR \approx 0,48=e^{-0,731}$). Estes resultados apontam para um aumento da incidência de padrões Benford atípicos quando os *accruals* discricionários são mais elevados, enquanto o imposto do período mais alto se relaciona com menor incidência. Trata-se de um sinal de uma possível existência de manipulação de resultados nesse quartil, ainda que circunscrito a 2023, Q3.

Concluindo, a evidência vai ao encontro do objetivo de que a combinação da Lei de Benford com o modelo de *accruals* discricionários de Jones permite uma melhor classificação das empresas com indícios de manipulação de resultados motivada pela minimização do imposto.

3.6. Discussão dos Resultados

Esta investigação aplicou, em sequência, a Lei de Benford (indicador LBC), o modelo de Jones (DAC e DACA), correlações de Pearson e Spearman, e modelos *logit* estimados por ano e por quartil definido com base em DAC. O objetivo foi avaliar indícios de manipulação de resultados com vista à minimização da carga fiscal no setor da restauração e similares. Esta combinação de métodos permite observar, primeiro, a existência de anomalias numéricas e de *accruals* discricionários, depois a

associação estatística entre LBC, DACA, NDIM e o ISRI, e, por fim, a probabilidade de ocorrência de anomalia Benford quando as variáveis relevantes são consideradas em conjunto.

Do ponto de vista descritivo-associativo, as correlações revelam um padrão predominante, com variações entre anos e quartis. A relação entre NDIM e ISRI é positiva, o que traduz um efeito de escala em que empresas maiores registam montantes de imposto mais elevados. Ao passo que, o LBC surge pouco relacionado com ISRI, NDIM e DACA. A leitura de Spearman, mais sensível à ordenação, evidencia, contudo, um traço robusto, em quartis extremos de DAC (Q1 e Q4) e em vários anos, as variáveis DACA e ISRI tomam um valor negativo, ao passo que NDIM–ISRI mantém uma associação positiva predominante, compatível com um efeito de escala (empresas maiores pagam mais imposto). Em 2023 Q3, surgem correlações estatisticamente significativas que não configuram um padrão temporal contínuo, mas revelam dinâmicas localizadas: o LBC relaciona-se positivamente com DACA e negativamente com NDIM, sugerindo indícios pontuais de manipulação de resultados nesse período. No conjunto de 2020 a 2023, porém, não há evidência de associação sistemática entre LBC e ISRI/NDIM/DACA.

Os modelos *logit* aprofundam esta leitura. Em diversos anos e quartis, com particular nitidez em 2022 e 2023, o coeficiente de ISRI é negativo e estatisticamente significativo. Isto indica que as observações com LBC apresentam menor imposto do período, mesmo mantendo constantes as variáveis DACA e NDIM. Este resultado está alinhado com o objetivo do estudo e reforça a plausibilidade de minimização fiscal associada a registos com sinais de anomalia numérica. O efeito de DACA é, em regra, não significativo. Existe, porém, um foco localizado em 2023 no quadril 3, onde DACA surge positivo e significativo no *logit*, o que aponta para maior probabilidade de a observação ser classificada como de LBC, quando os *accruals* discricionários aumentam. Trata-se de um sinal localizado, sem carácter generalizado. A NDIM, quando condicionada por quartil, não altera de forma consistente a probabilidade da LBC, o que sugere que a dimensão explica essencialmente a escala fiscal. Este padrão constitui um indício de possível manipulação de resultados nesse quartil, ainda que circunscrito a 2023 Q3.

As contagens de LBC mostram um padrão temporal estável entre 2020 e 2023, com proporções anuais semelhantes de observações sinalizadas e um conjunto de códigos críticos que se repete ao longo do tempo. Quando se estratifica por dimensão, o maior volume absoluto de ocorrências concentra-se nas microempresas. Esta concentração é compatível com a maior representatividade deste grupo na amostra e, por si só, não implica maior propensão.

Integrando as duas peças de evidência, correlações e modelo *logit*, conclui-se que não existe prova de um comportamento generalizado de manipulação de resultados no setor. Há, no entanto, um indício concreto e localizado em 2023 Q3, que é o alinhamento positivo entre DACA e LBC e o sinal negativo de ISRI face ao LBC sugerem manipulação de resultados nesse quadril e por sua vez, práticas

localizadas de minimização fiscal. Estes indícios emergem nas caudas da distribuição de *accruals*, onde os DACA e os ISRI são frequentemente negativos, e num subconjunto de observações de 2023-Q3, onde DACA se alinha positivamente com LBC enquanto ISRI mantém sinal negativo. Estes sinais sugerem que parte das entidades poderá gerir *accruals* em momentos específicos, com reflexo no imposto do período. Contudo, trata-se de associações estatísticas que constituem prova de manipulação intencional ou evasão. A inferência causal é demonstrável com os métodos utilizados e os padrões observados para o Q3.

Em termos de implicações, recomenda-se priorizar auditorias analíticas em anos e quartis onde as evidências descritivas apontam para as variáveis DACA e ISRI negativas e no caso particular do ano de 2023 no Q3, bem como complementar a análise com métricas fiscais mais detalhadas (ETR, impostos a caixa (*cash taxes*)) e com a reconciliação de diferenças temporárias e permanentes.

O objetivo do estudo foi atingido de forma prudente. A evidência aponta para indícios de minimização fiscal, não generalizados, mas concentrados em segmentos e períodos específicos, com destaque para os anos de 2023 e particular o terceiro quadril. Estes indícios são suficientes para orientar a monitorização e auditoria analítica. No seu conjunto, a evidência não sustenta a existência de evasão fiscal generalizada e não permite concluir que um grupo tem maior propensão apenas por apresentar mais casos em termos absolutos. Todavia, em 2023 Q3, observa-se um indício localizado de manipulação de resultados, refletido no alinhamento positivo entre DACA e LBC e no sinal negativo de ISRI face ao LBC.

Considerando o universo da amostra e a totalidade dos anos, a evidência empírica não sustenta a H1, conseqüentemente, a hipótese é rejeitada em termos globais. Todavia, na subamostra de 2023 (Q3) obtêm-se evidências estatisticamente significativas e consistentes com a H1, pelo que a hipótese é validada para esse ano e quartil. Concluindo, entre as empresas classificadas no Q3 em 2023 (predominantemente de micro dimensão), a combinação da Lei de Benford com o modelo de Jones (1991) apresenta maior capacidade classificativa na identificação de entidades com indícios de manipulação de resultados visando a minimização do imposto.

Conclusões, Limitações e Linhas de Investigação Futuras

Neste estudo, procurou-se avaliar a eficácia da combinação da Lei de Benford com o modelo baseado nos *accruals* discricionários - Modelo Jones (1991) - para aferir a QIF e identificar indícios de possível manipulação dos resultados de empresas do setor da restauração e similares e o impacto que isto tem na minimização do pagamento de impostos ao Estado.

Os resultados obtidos permitem concluir que não há evidência de manipulação de resultados generalizada no setor da restauração e similares. Identificam-se, porém, indícios localizados de minimização fiscal em determinados anos e quartis, em especial no terceiro quartil de 2023, onde se observa um sinal localizado de manipulação de resultados. O modelo logit indica que, nesse período, *accruals* discricionários mais elevados (DACA) aumentam a probabilidade de ocorrência de anomalias numéricas (LBC), enquanto níveis superiores de imposto (ISRI) reduzem essa probabilidade. Este padrão é compatível com práticas de manipulação de resultados orientadas à minimização da carga fiscal em contextos específicos, embora a evidência obtida, no seu conjunto, não constitua prova direta de evasão.

Contudo, a análise apresenta limitações que importa reconhecer. A inferência causal não é demonstrável com os métodos utilizados, baseados em evidência estatística e associativa. Além disso, a reduzida variabilidade de certas variáveis (nomeadamente NDIM em alguns quartis) limita a robustez das estimativas. Fatores externos, como alterações de políticas contabilísticas, sazonalidade, efeitos de calendário ou choques setoriais, podem igualmente influenciar os resultados, pelo que as conclusões devem ser interpretadas com prudência.

Em termos de implicações, recomenda-se priorizar auditorias analíticas nos anos e quartis, em que as evidências descritivas apontam para valores negativos das variáveis DACA e ISRI, com destaque para o ano de 2023, quartil 3. Deve igualmente ser complementada a análise com métricas fiscais, como a ETR e os impostos a caixa (*cash taxes*), bem como com a reconciliação de diferenças temporárias e permanentes. É ainda aconselhável aprofundar a investigação por subsetor/CAE e por porte empresarial, procedendo à realização de testes de robustez, de forma a reforçar a consistência empírica dos resultados.

Importa salientar, que uma das principais dificuldades deste estudo resultou da escassez de trabalhos anteriores que combinassem, de forma integrada, a aplicação da Lei de Benford e do modelo de Jones, o que limitou a consolidação teórica entre ambas as abordagens. A combinação de técnicas estatísticas com evidência documental ou procedimental poderá, no futuro, permitir comprovar com maior rigor a existência de práticas de minimização fiscal ou manipulação contabilística, no conjunto.

Referências

Achim, A. M., & Chis, A. O. (2014). Financial accounting quality and its defining characteristics. *Practical Application of Science* 2(5), 93-98. <https://www.semanticscholar.org/paper/FINANCIAL-ACCOUNTING-QUALITY-AND-ITS-DEFINING-Achim-Chi%C8%99/193774cc5bade744554678b6067a7fc894f8d458>

Arens, A. A., Elder, R. J., Beasley, M. S. & Chris, E. H. (2017). Auditing and assurance services (16.^a ed.). Pearson Education Limited. https://digilib.stekom.ac.id/assets/dokumen/ebook/feb_44bac1dd499213de626e2f232c01e8542ffef3bc_1652001111.pdf

Aloui, S., & Belanès, A. (2024). Assessment of financial information quality: Evidence from tunisian context. *Account and Financial Management Journal*, 9(2), 3277-3283. <https://doi.org/10.47191/afmj/v9i2.03>

Amorim, J., Sousa, B., Pereira, A., & Vaz, A. (2020). Tax incentives for companies to invest: The portuguese case in 2019. *Internacional Journal of Social Sciences*, 6(2), 94-110. <https://doi.org/10.20319/pijss.2020.62.94110>

Antunes, A. M., Teixeira, D. & Sousa, F. (23 de junho de 2023). *Lei de Benford: O braço esquerdo na deteção de fraude* [Comunicação e conferência]. Iberian Conference on Information Systems and Technologies e do Instituto Politécnico de Coimbra, Aveiro, Portugal. <https://research.ebsco.com/c/2ed5h2/viewer/pdf/pne5m3ixar>

Budianto, R., Susanto, D., Ginanjar, S., & Suyono E., (2024). Factors affecting on the quality of financial reports in small and medium size enterprises. *Risk Governance & Control: Financial Markets & Institutions*, 14(2), 15-24. <https://doi.org/10.22495/rgcv14i2p2>

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 78(4), 551-572. <https://www.jstor.org/stable/984802?origin=JSTOR-pdf>

Christensen, T. E., Huffman, A., Lewis-Occidental, M. F. & Valentim, K. (2022). A simple approach to better distinguish real earnings manipulation from strategy changes. *Contemporary Accounting Research*, 40(1), 406-450. <https://doi.org/10.1111/1911-3846.12830>

Comissão de Normalização Contabilística. (2021). *Sistema de Normalização Contabilística: Estrutura Conceptual*. CNC. https://www.cnc.min-financas.pt/_siteantigo/snc_projecto/snc_ec.pdf

Costa, C. M., & Soares, J. M. (2022). Standard jones and modified jones: An earnings management tutorial. *Revista de Administração Contemporânea*, 26(2), 1-13. <https://doi.org/10.1590/1982-7849rac2022200305.en>

Crispim, A. I. R. (2023). *Aplicação da lei de Benford na deteção de fraude ao nível das demonstrações financeiras: O caso do setor do turismo*. [Dissertação de mestrado, Instituto Superior de Economia e Gestão]. Repositório da Universidade de Lisboa. <https://repositorio.ulisboa.pt/bitstream/10400.5/29861/1/DM-AIRC-2023.pdf#page=17.72>

Cupertino, C. M. & Martinez, A. L. (2008). Qualidade da auditoria e earnings management: Risk assessment através do nível dos accruals discricionários. *Contabilidade Vista & Revista*, 19(3), 69-93. <https://www.redalyc.org/pdf/1970/197014588004.pdf>

Dinu, E. L. (2022). The concept of quality of financial accounting information and the relationship with IFRS standards: Analysis of the situation of companies listed on BVB. *Sciendo*, 16(1), 1353-1367. <https://doi.org/10.2478/picbe-2022-0123>

Dănescu, T., Spătăcean, I.-O., Popa, M.-A., & Sîrbu, C.-G. (2021). The Impact of Corporate Governance Mechanism over Financial Performance: Evidence from Romania. *Sustainability*, 13(19), 10494. <https://doi.org/10.3390/su131910494>

Dănescu, T., & Stejerean, R. M. (2022). Companies' behavior in measuring the quality of financial reports: Pre- and post-pandemic research. *Frontiers in psychology*. 13, 1-11. <https://doi.org/10.3389/fpsyg.2022.1005941>

Elisabeta, B. D., Lucian, C., Teodora, M. I. & Victoria, B. (2014). Accounting manipulation: An empirical study regarding managers' behaviour. *Annals of Faculty of Economics*, 1, 554-562. <https://anale.steconomiceuoradea.ro/volume/2014/n1/058.pdf>

Gillette, D. B. & Stinson, S. R. (2022). Who gets to play dirty? Using legitimacy theory to examine investor reactions to differing modes of corporate tax minimization. *Contemporary Accounting Research*, 39(4), 2596-2621. <https://doi.org/10.1111/1911-3846.12805>

Gorenc, M. (2019). Benford's law as a useful tool to determine fraud in financial statements. *Management*, 14(1), 19-31. 10.26493/1854-4231.14.19-31. 10.26493/1854-4231.14.19-31

Jara, E. G., Ebrero A.C., & Zapata R. E. (2011). Effect of international financial reporting standards on financial information quality. *Journal of Financial Reporting and Accounting*, 9(2), 176-196. <https://doi.org/10.1108/19852511111173121>

Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29(2), 193-228. <https://doi.org/10.2307/2491047>

Latorre, M. R. D. O. & Cardoso, M. R. A. (2001). Análise de séries temporais em epidemiologia: Uma introdução sobre os aspectos metodológicos. *Revista Brasileira De Epidemiologia*, 4(3), 145-152. <https://doi.org/10.1590/S1415-790X2001000300002>

Lesperance, M., Reed, W. J., Stephens, M. A., Tsao, C. & Wilton, B. (2016). Assessing

conformance with Benford's law: Goodness-of-fit tests and simultaneous confidence intervals. *Plos One*, 11(3), 1-20. <https://doi.org/10.1371/journal.pone.0151235>

Licerán-Gutiérrez, A., & Cano-Rodríguez, M. (2019). A review on the multidimensional analysis of earnings quality. *Revista de Contabilidad*, 22(1), 41-60. <https://doi.org/10.6018/rc-sar.22.1.354301>

Marques, B. A., Leal, D., & Rody, P. H. A. (2016). Contribuição da lei do bem para o planeamento tributário de uma unidade empresarial e a percepção dos contadores do estado do espírito santo sobre incentivos fiscais e planeamento tributário. *Revista Ambiente Contábil*, 8(2), 40-58. <https://doi.org/10.21680/2176-9036.2016v8n2id7789>

Martins, A. (2023). *Aplicação da lei de Benford na deteção de Indícios de subfaturação em empresas do setor de alojamento e restauração e similares no período de Pandemia Covid-19* [Dissertação de mestrado, Instituto Politécnico de Bragança]. Biblioteca Digital do IPB. <http://hdl.handle.net/10198/29266>

Mendonça, L. S., Rosa, B. R. & Leal, G. S. (2022). A utilização de inteligência artificial -machine learning e business intelligence- na deteção e prevenção de fraudes contábeis: Primeiras aproximações. *Enciclopédia Biosfera*, 19(41), 1-16. https://doi.org/10.18677/EnciBio_2022C1

Nigrini, M. (2011). *Forensic analytics: Methods and techniques for forensic accounting investigations*. (2.^a ed.) John Wiley & Sons, Inc. <https://books.google.pt/books?id=hSrfDwAAQBAJ&lpg=PP1&hl=pt-PT&pg=PP1#v=onepage&q&f=false>

Nigrini, M. J. & Mittermaier, L. J. (1997). The use of Benford's law as an aid in analytical Procedures. *A Journal of Practice & Theory*, 16(2), 52-67. <https://research.ebsco.com/c/z72xap/search/details/nylk7564lv?db=bth>

Nuhu, M. S., Ahmad, Z., & Zhee, L. Y. (2024). Nexus among disclosure quality, discretionary accruals and real earnings management practices: An empirical analysis of Malaysian public firms. *Journal of Corporate Accounting & Finance*, 35(4), 121-138. <https://doi.org/10.1002/jcaf.22720>

Oliveira, C. D., De Caroli, A. A., Amaral A. S. & Vilca, O. L. (2014). Deteção de fraudes, anomalias e erros em análise de dados contábeis: Um estudo com base em outliers. *Revista Eletrónica do Departamento de Ciências Contábeis & Departamento de Atuária e Métodos Quantitativos de Fea*, 1(1), 102-127. <https://doi.org/10.23925/2446-9513.2014V1I1P102-127>

Pais, C. & Dias, C. A. (2022). The implications of book-tax conformity and tax change for the earnings management of portuguese micro firms. *Journal of International Accounting, Auditing and Taxation* 46. 1-13. <https://doi.org/10.1016/j.intaccaudtax.2022.100448>

Renkas, J., Goncharenko, O. & Lukianets, O. (2016). Quality of financial reporting: Approaches to measuring. *International Journal of Accounting and Economics Studies*, 4(1), 1-5. <https://doi.org/10.14419/ijaes.v4i1.5509>

Silva, B. C. O. & Nóbrega, R. S. (2021). Dados discrepantes ou outliers: Avaliação da quadra chuvosa do semiárido do m, brasil. *Entre-Lugar*, 13(24), 346-366. <https://doi.org/10.30612/rel.v12i24.15188>

Silva, D. & Cerqueira, A. (2021). Financial Reporting Quality and Investors' Divergence of Opinion. *Accounting Perspectives*, 20(1), 1-164. <https://doi.org/10.1111%2F1911-3838.12250>

Silva, T. S. T., Nogueira, F. M. D., & Maia, A. J. M. (2022). Motivações na esfera do (in)cumprimento fiscal. *Brazilian Journal of Development*, 8(4), 27374-27407. <https://doi.org/10.34117/bjdv8n4-306>

Silva, W., Travassos, S., & Costa, J. (2017). Using the newcomb-benford law as a deviation identification method in continuous auditing environments: A proposal for detecting deviations over time. *Revista Contabilidade & Finanças*, 28(73), 11-26. <https://doi.org/10.1590/1808-057x201702690>

Souza, A. A. de, Fonseca, S. E., & Rezende, L. D. (2024). Análise comparativa da qualidade das informações contábeis entre empresas francesas, brasileiras e latino americanas. *Revista Ambiente Contábil*, 16(2), 24–43. <https://doi.org/10.21680/2176-9036.2024v16n2ID32914>

Swanson, D., Cho, M., & Eltinge, J. (2003). Detecting possibly fraudulent or error-prone survey data using Benford's law. *Bureau of Labor Statistics*, 4172-4177.

Szumilas, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 19(3), 227-229. https://pmc.ncbi.nlm.nih.gov/articles/PMC2938757/pdf/ccap19_3p227.pdf

Trussel, J. (2003). Assessing potential accounting manipulation: The financial characteristics of charitable organizations with higher than expected program-spending. *SAGE Publications*, 32(4), 493-661. <https://doi.org/10.1177/0899764003257459>

Valaskova, K. & Fedorko, R. (2021). Beneish M-score: A measure of fraudulent financial transactions in global environment? *SHS Web of Conferences*, 92. 1-10. <https://doi.org/10.1051/shsconf/20219202064>

Yan S. (2024). Corporate Social Responsibility and Financial Information Quality. *Journal of Business and Economic Studies*, 28(1), 99-111. <https://doi.org/10.53462/vol28i1a5>

Decreto-Lei n.º 162/2014, 211 (I Série) Diário da República (Portugal), 1-27 (31 de outubro de 2014). <https://diariodarepublica.pt/dr/legislacao-consolidada/decreto-lei/2014-59423292>

Apêndice

Apêndice 1 – Dimensão da amostra total e dos LBC=1, por dimensão da empresa

NDIM	N total	N LBC=1
1	87364	15122
2	1715	86
3	106	17
4	34	8

Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

Apêndice 2 – Dimensão da amostra por ano

Ano	N total
2020	22420
2021	22714
2022	22727
2023	21356
2024	2

Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

Apêndice 3 - Tabela de estatística descritiva

Descrição	ANO	A	AFT	TPC	FOC	VPS	GRDA	RAI	ISRI	EBITDA	PRISR	ACC
Média	2020	245 227,30 €	96 995,12 €	97 616,51 €	20 124,69 €	236 625,10 €	12 323,08 €	-16 127,43 €	1 150,45 €	-2 469,67 €	-252,95 €	-1 136,20 €
	2021	268 113,90 €	100 498,30 €	109 540,80 €	20 376,34 €	284 932,60 €	12 606,18 €	6 514,45 €	3 107,44 €	20 604,53 €	11 987,01 €	-20 388,92 €
	2022	293 696,40 €	107 465,40 €	122 632,10 €	19 793,36 €	431 521,80 €	13 528,91 €	23 198,19 €	5 128,44 €	38 389,13 €	-4 960,04 €	-18 883,51 €
	2023	324 798,10 €	119 461,70 €	131 048,20 €	20 174,40 €	502 237,70 €	15 277,94 €	28 153,36 €	6 477,03 €	46 683,96 €	-23 310,26 €	-13 357,60 €
	2024	1 173 481,00 €	410 962,90 €	684 515,90 €	0,00 €	3 016 646,00 €	101 927,84 €	26 536,25 €	5 830,89 €	126 646,33 €	NA	-53 812,41 €
Desvio Padrão	2020	2 158 958,00 €	1 293 289,00 €	633 481,90 €	283 801,60 €	1 608 210,00 €	151 227,23 €	192 505,18 €	41 615,27 €	256 529,07 €	306 557,37 €	221 797,70 €
	2021	2 283 549,00 €	1 369 946,00 €	686 078,30 €	203 017,40 €	1 876 159,00 €	152 162,87 €	258 760,31 €	66 378,13 €	363 891,98 €	413 547,18 €	219 837,60 €
	2022	2 465 938,00 €	1 456 701,00 €	763 717,30 €	148 116,60 €	2 559 169,00 €	161 871,33 €	389 986,01 €	103 638,44 €	499 160,84 €	337 901,45 €	257 158,80 €
	2023	3 560 729,00 €	1 723 106,00 €	868 592,60 €	206 455,60 €	3 151 735,00 €	230 131,18 €	548 777,41 €	162 061,15 €	669 056,70 €	481 075,90 €	328 079,20 €
	2024	1 583 894,00 €	517 981,30 €	898 880,30 €	0,00 €	4 168 160,00 €	137 541,32 €	35 949,33 €	7 579,58 €	166 616,42 €	NA	13 661,45 €
Mediana	2020	66 135,62 €	12 874,18 €	29 369,97 €	0,00 €	95 663,68 €	2 180,49 €	-9 302,58 €	0,00 €	-5 829,83 €	0,00 €	-1 768,33 €
	2021	72 774,09 €	13 442,76 €	34 269,70 €	0,00 €	114 364,40 €	2 277,71 €	-301,84 €	132,53 €	1 937,68 €	0,00 €	-7 309,65 €
	2022	80 307,30 €	14 121,94 €	39 073,69 €	0,00 €	173 765,30 €	2 469,99 €	2 226,46 €	332,28 €	6 134,86 €	0,00 €	-6 217,30 €
	2023	86 342,90 €	15 726,08 €	42 226,01 €	0,00 €	202 820,70 €	2 792,98 €	2 687,58 €	464,72 €	7 293,95 €	0,00 €	-5 108,11 €
	2024	1 173 481,00 €	410 962,90 €	684 515,90 €	0,00 €	3 016 646,00 €	101 927,84 €	26 536,25 €	5 830,89 €	126 646,33 €	NA	-53 812,41 €
Mínimo	2020	4,88 €	0,00 €	0,00 €	0,00 €	1,63 €	-35 490,79 €	-6 950 486,65 €	-2 063 357,27 €	-5 207 299,20 €	-2 036 829,22 €	-20 997 760,00 €
	2021	8,48 €	0,00 €	0,00 €	0,00 €	4,40 €	-14 343,33 €	-4 993 023,11 €	-1 326 208,31 €	-4 321 906,07 €	-1 217 376,48 €	-16 640 870,00 €
	2022	1,71 €	0,00 €	0,00 €	0,00 €	2,15 €	-1 454,72 €	-5 108 887,41 €	-5 428 117,17 €	-1 856 132,11 €	-3 640 194,88 €	-21 602 600,00 €
	2023	4,94 €	0,00 €	0,00 €	0,00 €	1,88 €	-21 065,94 €	-7 853 006,91 €	-5 772 701,78 €	-2 245 510,73 €	-5 455 638,52 €	-33 277 640,00 €
	2024	53 498,54 €	44 694,79 €	48 911,54 €	0,00 €	69 312,30 €	4 671,44 €	1 116,23 €	471,32 €	8 830,73 €	NA	-63 472,52 €
Máximo	2020	190 580 900,00 €	137 322 000,00 €	50 801 900,00 €	31 404 160,00 €	102 840 000,00 €	13 433 797,69 €	19 869 920,64 €	4 854 026,54 €	28 717 492,90 €	8 266 504,47 €	3 559 552,00 €
	2021	200 924 700,00 €	148 882 000,00 €	44 117 390,00 €	19 385 900,00 €	112 282 100,00 €	13 595 662,48 €	31 348 945,30 €	8 697 526,20 €	40 535 384,89 €	11 171 371,21 €	4 145 710,00 €
	2022	216 166 100,00 €	159 890 900,00 €	50 489 770,00 €	11 808 300,00 €	143 751 300,00 €	14 983 551,94 €	46 949 729,97 €	12 957 576,81 €	56 067 007,89 €	8 265 895,89 €	12 959 630,00 €
	2023	355 019 600,00 €	185 635 000,00 €	53 846 910,00 €	18 184 670,00 €	199 118 300,00 €	27 341 389,85 €	65 331 558,49 €	19 620 588,06 €	74 595 887,23 €	10 686 884,38 €	17 036 360,00 €
	2024	2 293 463,00 €	777 231,00 €	1 320 120,00 €	0,00 €	5 963 980,00 €	199 184,24 €	51 956,26 €	11 190,46 €	244 461,93 €	NA	-44 152,31 €
Coeficiente de Variação	2020	880,4%	1333,4%	648,9%	1410,2%	679,6%	1227,2%	-1193,7%	3617,3%	-10387,2%	-121194,2%	-19521,0%
	2021	851,7%	1363,2%	626,3%	996,3%	658,5%	1207,0%	3972,1%	2136,1%	1766,1%	3450,0%	-1078,2%
	2022	839,6%	1355,5%	622,8%	748,3%	593,1%	1196,5%	1681,1%	2020,9%	1300,3%	-6812,5%	-1361,8%
	2023	1096,3%	1442,4%	662,8%	1023,4%	627,5%	1506,3%	1949,2%	2502,1%	1433,2%	-2063,8%	-2456,1%
	2024	135,0%	126,0%	131,3%	NA	138,2%	134,9%	135,5%	130,0%	131,6%	NA	-25,4%

Apêndice 3 - Tabela de estatística descritiva (Continuação)

Descrição	ANO	VPS_1	ACCD	INVA	VEPSD	AFTD	DAC	LBC	NDIM	DACA	ISRED
Média	2020	-128 229,30 €	-0,16 €	0,00 €	-1,57 €	0,35 €	0,00 €	0,17 €	1,01 €	2,78 €	0,01 €
	2021	42 699,59 €	-0,47 €	0,00 €	1,44 €	0,36 €	0,00 €	0,17 €	1,02 €	1,16 €	0,01 €
	2022	148 529,20 €	-0,42 €	0,00 €	1,94 €	0,37 €	0,00 €	0,18 €	1,03 €	0,87 €	0,02 €
	2023	70 561,95 €	-0,59 €	0,00 €	1,16 €	0,36 €	0,00 €	0,17 €	1,04 €	1,32 €	0,04 €
	2024	133 148,50 €	-0,37 €	0,00 €	-0,50 €	0,55 €	0,00 €	0,50 €	1,50 €	0,00 €	0,01 €
Desvio Padrão	2020	747 362,60 €	36,69 €	0,00 €	53,76 €	0,77 €	20,32 €	0,38 €	0,12 €	20,12 €	0,19 €
	2021	286 193,90 €	11,62 €	0,00 €	102,63 €	0,87 €	9,35 €	0,37 €	0,14 €	9,28 €	0,06 €
	2022	806 124,60 €	6,73 €	0,00 €	15,34 €	1,14 €	5,06 €	0,38 €	0,19 €	4,98 €	0,09 €
	2023	663 044,80 €	33,69 €	0,00 €	26,35 €	0,86 €	18,74 €	0,38 €	0,20 €	18,70 €	2,41 €
	2024	291 377,20 €	0,48 €	0,00 €	0,94 €	0,23 €	0,00 €	0,71 €	0,71 €	0,00 €	0,00 €
Mediana	2020	-42 801,13 €	-0,03 €	0,00 €	-0,68 €	0,23 €	-1,15 €	0,00 €	1,00 €	1,44 €	0,00 €
	2021	9 646,07 €	-0,10 €	0,00 €	0,13 €	0,23 €	-0,22 €	0,00 €	1,00 €	0,46 €	0,00 €
	2022	52 735,31 €	-0,08 €	0,00 €	0,72 €	0,22 €	-0,04 €	0,00 €	1,00 €	0,28 €	0,00 €
	2023	20 701,75 €	-0,07 €	0,00 €	0,24 €	0,23 €	-0,24 €	0,00 €	1,00 €	0,49 €	0,01 €
	2024	133 148,50 €	-0,37 €	0,00 €	-0,50 €	0,55 €	0,00 €	0,50 €	1,50 €	0,00 €	0,01 €
Mínimo	2020	-45 159 830,00 €	-3 768,02 €	0,00 €	-7 073,39 €	0,00 €	-2 167,82 €	0,00 €	1,00 €	0,00 €	-0,44 €
	2021	-10 994 800,00 €	-1 421,33 €	0,00 €	-230,32 €	0,00 €	-1 152,34 €	0,00 €	1,00 €	0,00 €	-0,64 €
	2022	-2 359 757,00 €	-527,29 €	0,00 €	-108,26 €	0,00 €	-244,85 €	0,00 €	1,00 €	0,00 €	-0,27 €
	2023	-5 732 145,00 €	-4 104,94 €	0,00 €	-674,14 €	0,00 €	-1 377,49 €	0,00 €	1,00 €	0,00 €	-0,42 €
	2024	-72 886,28 €	-0,71 €	0,00 €	-1,17 €	0,38 €	0,00 €	0,00 €	1,00 €	0,00 €	0,00 €
Máximo	2020	3 007 407,00 €	3 635,24 €	0,10 €	257,36 €	65,65 €	1 164,50 €	1,00 €	4,00 €	2 167,82 €	24,48 €
	2021	21 181 090,00 €	218,59 €	0,02 €	15 373,98 €	93,77 €	373,18 €	1,00 €	4,00 €	1 152,34 €	4,32 €
	2022	45 551 290,00 €	117,33 €	0,05 €	1 220,40 €	96,69 €	228,22 €	1,00 €	4,00 €	244,85 €	6,96 €
	2023	60 869 930,00 €	1 415,60 €	0,41 €	2 937,08 €	58,88 €	1 984,72 €	1,00 €	4,00 €	1 984,72 €	351,72 €
	2024	339 183,30 €	-0,03 €	0,00 €	0,17 €	0,71 €	0,00 €	1,00 €	2,00 €	0,00 €	0,01 €
Coeficiente de Variação	2020	-582,8%	-22387,0%	1598,6%	-3416,5%	219,7%	222456200000000000%	221,1%	11,9%	722,6%	1721,7%
	2021	670,2%	-2481,3%	581,2%	7122,5%	243,9%	-433896400000000000%	222,4%	13,4%	800,1%	439,0%
	2022	542,7%	-1621,1%	1060,6%	791,8%	303,8%	301142100000000000%	216,8%	18,0%	573,2%	467,3%
	2023	939,7%	-5743,8%	3533,5%	2277,1%	236,2%	1459013000000000000%	221,5%	19,5%	1414,0%	6370,3%
	2024	218,8%	-129,4%	133,0%	-188,7%	42,8%	141,4%	141,4%	47,1%	141,4%	40,6%

Fonte: Elaboração própria, com recurso ao Excel e ao software RStudio.

Apêndice 4 - Frequências do primeiro e segundo dígito

Números	Contagem dos números	Frequência observada	Frequência teórica	Desvio	P-Value teste Z	P-Value teste χ^2	Nível de significância
10	5371	4,23%	4,14%	0,001	0,109	0,117	>=0,05
11	4929	3,88%	3,78%	0,001	0,057	0,062	>=0,05
12	4593	3,62%	3,48%	0,001	0,006	0,007	<0,05
13	4135	3,26%	3,22%	0,000	0,453	0,460	>=0,05
14	3898	3,07%	3,00%	0,001	0,128	0,134	>=0,05
15	3789	2,98%	2,80%	0,002	0,000	0,000	<0,05
16	3304	2,60%	2,63%	0,000	0,483	0,489	>=0,05
17	3261	2,57%	2,48%	0,001	0,051	0,054	>=0,05
18	3126	2,46%	2,35%	0,001	0,008	0,009	<0,05
19	2925	2,30%	2,23%	0,001	0,069	0,072	>=0,05
20	2775	2,18%	2,12%	0,001	0,103	0,106	>=0,05
21	2584	2,03%	2,02%	0,000	0,720	0,723	>=0,05
22	2530	1,99%	1,93%	0,001	0,111	0,115	>=0,05
23	2436	1,92%	1,85%	0,001	0,065	0,068	>=0,05
24	2285	1,80%	1,77%	0,000	0,479	0,483	>=0,05
25	2207	1,74%	1,70%	0,000	0,344	0,349	>=0,05
26	2095	1,65%	1,64%	0,000	0,770	0,771	>=0,05
27	2016	1,59%	1,58%	0,000	0,822	0,824	>=0,05
28	1911	1,50%	1,52%	0,000	0,573	0,576	>=0,05
29	1792	1,41%	1,47%	-0,001	0,069	0,071	>=0,05
30	1878	1,48%	1,42%	0,001	0,101	0,103	>=0,05
31	1691	1,33%	1,38%	0,000	0,147	0,150	>=0,05
32	1659	1,31%	1,34%	0,000	0,349	0,352	>=0,05
33	1614	1,27%	1,30%	0,000	0,418	0,421	>=0,05
34	1538	1,21%	1,26%	0,000	0,125	0,127	>=0,05
35	1527	1,20%	1,22%	0,000	0,492	0,495	>=0,05
36	1517	1,19%	1,19%	0,000	0,883	0,884	>=0,05
37	1431	1,13%	1,16%	0,000	0,294	0,297	>=0,05
38	1371	1,08%	1,13%	0,000	0,101	0,103	>=0,05
39	1349	1,06%	1,10%	0,000	0,201	0,203	>=0,05
40	1403	1,10%	1,07%	0,000	0,264	0,267	>=0,05
41	1264	1,00%	1,05%	-0,001	0,072	0,074	>=0,05
42	1284	1,01%	1,02%	0,000	0,697	0,699	>=0,05
43	1214	0,96%	1,00%	0,000	0,127	0,129	>=0,05
44	1150	0,91%	0,98%	-0,001	0,011	0,011	<0,05
45	1190	0,94%	0,95%	0,000	0,519	0,521	>=0,05
46	1107	0,87%	0,93%	-0,001	0,021	0,021	<0,05
47	1098	0,86%	0,91%	0,000	0,062	0,063	>=0,05
48	1158	0,91%	0,90%	0,000	0,539	0,540	>=0,05
49	1042	0,82%	0,88%	-0,001	0,029	0,030	<0,05
50	1085	0,85%	0,86%	0,000	0,824	0,825	>=0,05
51	974	0,77%	0,84%	-0,001	0,003	0,003	<0,05
52	1010	0,80%	0,83%	0,000	0,207	0,209	>=0,05
53	987	0,78%	0,81%	0,000	0,168	0,170	>=0,05
54	996	0,78%	0,80%	0,000	0,611	0,612	>=0,05
55	1005	0,79%	0,78%	0,000	0,724	0,725	>=0,05
56	926	0,73%	0,77%	0,000	0,106	0,107	>=0,05
57	934	0,74%	0,76%	0,000	0,412	0,414	>=0,05
58	931	0,73%	0,74%	0,000	0,697	0,698	>=0,05
59	856	0,67%	0,73%	-0,001	0,019	0,020	<0,05
60	1007	0,79%	0,72%	0,001	0,002	0,002	<0,05

Números	Contagem dos números	Frequência observada	Frequência teórica	Desvio	P-Value teste Z	P-Value teste χ^2	Nível de significância
61	863	0,68%	0,71%	0,000	0,256	0,257	>=0,05
62	800	0,63%	0,69%	-0,001	0,005	0,005	<0,05
63	850	0,67%	0,68%	0,000	0,525	0,526	>=0,05
64	794	0,63%	0,67%	0,000	0,036	0,036	<0,05
65	797	0,63%	0,66%	0,000	0,119	0,120	>=0,05
66	776	0,61%	0,65%	0,000	0,062	0,063	>=0,05
67	838	0,66%	0,64%	0,000	0,465	0,467	>=0,05
68	788	0,62%	0,63%	0,000	0,542	0,543	>=0,05
69	763	0,60%	0,62%	0,000	0,275	0,276	>=0,05
70	776	0,61%	0,62%	0,000	0,818	0,818	>=0,05
71	784	0,62%	0,61%	0,000	0,651	0,652	>=0,05
72	765	0,60%	0,60%	0,000	0,880	0,880	>=0,05
73	746	0,59%	0,59%	0,000	0,870	0,870	>=0,05
74	720	0,57%	0,58%	0,000	0,452	0,453	>=0,05
75	690	0,54%	0,58%	0,000	0,132	0,133	>=0,05
76	721	0,57%	0,57%	0,000	0,998	0,998	>=0,05
77	675	0,53%	0,56%	0,000	0,167	0,168	>=0,05
78	742	0,58%	0,55%	0,000	0,137	0,138	>=0,05
79	666	0,52%	0,55%	0,000	0,289	0,291	>=0,05
80	760	0,60%	0,54%	0,001	0,004	0,004	<0,05
81	658	0,52%	0,53%	0,000	0,468	0,470	>=0,05
82	659	0,52%	0,53%	0,000	0,709	0,710	>=0,05
83	648	0,51%	0,52%	0,000	0,623	0,624	>=0,05
84	715	0,56%	0,51%	0,000	0,015	0,015	<0,05
85	593	0,47%	0,51%	0,000	0,040	0,040	<0,05
86	601	0,47%	0,50%	0,000	0,145	0,146	>=0,05
87	578	0,46%	0,50%	0,000	0,036	0,037	<0,05
88	646	0,51%	0,49%	0,000	0,362	0,363	>=0,05
89	607	0,48%	0,49%	0,000	0,707	0,707	>=0,05
90	599	0,47%	0,48%	0,000	0,670	0,670	>=0,05
91	575	0,45%	0,47%	0,000	0,256	0,257	>=0,05
92	616	0,49%	0,47%	0,000	0,419	0,420	>=0,05
93	558	0,44%	0,46%	0,000	0,187	0,188	>=0,05
94	596	0,47%	0,46%	0,000	0,610	0,611	>=0,05
95	602	0,47%	0,45%	0,000	0,309	0,310	>=0,05
96	589	0,46%	0,45%	0,000	0,466	0,467	>=0,05
97	544	0,43%	0,45%	0,000	0,360	0,361	>=0,05
98	599	0,47%	0,44%	0,000	0,099	0,099	>=0,05
99	555	0,44%	0,44%	0,000	0,979	0,979	>=0,05

Fonte: Elaboração própria, com recurso ao Excel.

Apêndice 5 - Dimensão da amostra total, por LBC total e LBC=1 a sua correspondência em % em relação à LBC total e extremos dos quartis por ano

Ano	Quartil	N total	N LBC	Min. DAC	Max. DAC	N LBC=1	Prob. LBC=1
2020	Q1	5605	5605	-2167,815	-1,731	948	16,91%
	Q2	5605	5605	-1,730	-1,150	937	16,72%
	Q3	5605	5605	-1,150	-0,182	985	17,57%
	Q4	5605	5605	-0,181	1164,502	938	16,74%
2021	Q1	5679	5679	-1152,342	-0,589	941	16,57%
	Q2	5679	5679	-0,589	-0,221	939	16,53%
	Q3	5678	5678	-0,220	0,183	1008	17,75%
	Q4	5678	5678	0,184	373,181	931	16,40%
2022	Q1	5682	5682	-244,850	-0,275	977	17,19%
	Q2	5682	5682	-0,275	-0,037	997	17,55%
	Q3	5682	5682	-0,037	0,277	990	17,42%
	Q4	5681	5681	0,277	228,216	1024	18,03%
2023	Q1	5339	5339	-1377,494	-0,613	930	17,42%
	Q2	5339	5339	-0,613	-0,237	904	16,93%
	Q3	5339	5339	-0,237	0,262	891	16,69%
	Q4	5339	5339	0,262	1984,722	892	16,71%

Fonte: Elaboração própria, com recurso ao Excel.